


```
1090      1100      1110      1120      1130      1140      1150
      X
      10      20      X
      KLGFFRSARRRRPGLDTPKYLE
      |||||
      LLLLALLVIALRKLGFFRSARRRRPGLDTPKYLE
      1160      X 1170      1180      X
```


using sw model

2004, 12:52:17 ; Search time 1.2067 Seconds
(without alignments)
1913.143 Million cell updates/sec

4-2_COPY_1165_1188
RREFGLDPTPKVLE 24

Gapext 0.5

96191526 residues

ing chosen parameters: 283366

00000

ch 0%
ch 100%
st 45 summaries

```

PIR 78:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*

```

er of results predicted by chance to have a equal to the score of the result being printed, vis of the total score distribution.

SUMMARIES

DB	ID	Description
2	JS0685	glutamate receptor
2	A84358	trans lesion repair
2	E70338	probable aspartate
2	T41362	hypothetical prote
2	T38490	hypothetical prote
2	140485	surfactin syntheta
2	K83525	TolQ protein PA096
2	A10961	conserved hypothet
2	S78089	G-protein signalin
2	167945	3',5'-cyclic-nucle
2	167945	aspartate-tRNA lig
1	S42047	probable histidine
2	H95367	3',5'-cyclic-nucle
2	S71626	DNA polymerase typ
2	G84384	CAMP phosphodieste
2	B32904	Nck, Ash and phosph
2	PC4427	hypothetical prote
2	S77409	Protein Putative A
2	D96693	conserved hypothet
2	G37259	F44E2.3 protein -
2	S48822	hypothetical prote
2	A12449	hypothetical prote
2	E95416	hypothetical prote
2	C87019	hypothetical prote
2	JH0572	transcriptional regu
2	A48990	hypothetical prote
2	T25567	hypothetical prote
2	AF2427	hypothetical prote
2	G70639	hypothetical prote
2	JQ1383	hypothetical 84 k

30	44	35.2	1214	2	S28499	probable finger pr
31	44	35.2	1234	2	A34911	band 3-related pro
32	44	35.2	1237	2	A31789	band 3-related pro
33	44	35.2	1237	2	A56764	band 3-related pro
34	44	35.2	1450	2	A84780	probable ABC trans
35	44	35.2	1810	2	T31032	probable voltage-g
36	44	35.2	2649	2	T51023	hypothetical prote
37	43	34.4	340	2	C95940	probable allantoic
38	43	34.4	348	2	G83976	cytochrome caa3 ox
39	43	34.4	353	2	F75581	hypothetical prote
40	43	34.4	417	2	F95953	probable glycosylt
41	43	34.4	437	2	A72498	probable DNA/panto
42	43	34.4	453	2	G59070	coenzyme F390 synt
43	43	34.4	462	2	G87500	hypothetical prote
44	43	34.4	584	2	B53109	3',5'-cyclic-nucle
45	43	34.4	672	2	I61259	3',5'-cyclic-nucle

ALIGNMENTS

RESULT 1

JS0685
 glutamate receptor chain KA-1 precursor - rat
 N;Alternate names: kainate receptor kal
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 1-7-Jul-1992 #sequence_revision 17-Jul
 C;Accession: J05685; S18266; J19154
 R;Herb, A.; Burnashev, N.; Werner, P.; Sakmann
 Neuron 8, 775-785, 1992
 A;Title: The KA-2 subunit of excitatory amino
 A;Accession number: JH0592; MUID:92232306; PM
 A;Reference: JS0685
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-956 <HER>
 A;Cross-references: EMBL:X59996; NID:g57614;
 R;Werner, P.; Voigt, M.; Keinänen, K.; Wisden
 Nature 351, 742-744, 1991
 A;Title: Cloning of a putative high-affinity
 A;Accession number: S18266; MUID:91287798; PM
 A;Reference: JS8266
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-915, 916HRRPQVLPALPGAGSTVASAQR
 A;Cross-references: EMBL:X59996
 C;Superfamily: Glutamate receptor; Glutamate
 C;Keywords: neurotransmitter receptor; transmembrane
 F;1-20/Domain: signal sequence #status predicted
 F;21-956/Product: glutamate receptor chain KA-1
 F;420-938/Domain: glutamate receptor homology
 F;457-566/Domain: transmembrane #status predicted
 F;591-609/Domain: transmembrane #status predicted
 F;620-938/Domain: transmembrane #status predicted

Query Match	41.6%	Score 52	DB 2	Length 956
Best Local Similarity	68.8%	Pred.No. 11		
Matches 11	Conservative 0	Mismatches 5	Indels 0	Gaps 0
Cy	9	RRRREGLDPTPKVLE	24	
db	862	RRRSGLPPPPVLE	877	

2. T. J. S. A.

trans lesion repair [imported]. - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: A84258
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.
 R:Leithauer, B.; Kaller, R.; Cruz, R.; Danson, M.G.; Hough, D.W.; Maddocks

Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: A84258
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AE004437; NID:g10580569; PIDN:AAG19429.1; GSPDB:GN00138
C:Genetics:
A:Gene: yqjH

```

Query Match          39.2%; Score 49; DB 2; Length 447;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      4  FFSARRRRRPGLDPTFKVL 23
      | : | | | | | | :
Db      62  FYASCRRRREPALRGAPVV 81

RESULT 3
E70338
probable aspartate transaminase (EC 2.6.1.1) aspC3 [similarity] - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000
C;Accession: E70338
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.

```

V. Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: E70338
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-373 <AQP>
A:Cross-references: GB:AE000689; NID:g2983082; PIDN:AAC06689.1; PID:g2983083; GB:AE000689
A:Experimental source: strain VF5

C:genetics: C
A:Gene: aspC3
C:Superfamily: aspartate transaminase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F.222/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy	13	EPGLDPTPKVLE	24
		: : :	
D _b	32	EPDLEPSPKVM	43

RESULT 4
T41362
Hypothetical protein SPCC4G3.14 - fission yeast (*Schizosaccharomyces pombe*)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: T41362
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21918
A:Accession: T41362
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-528 <WOO>
A:Cross-references: EMBL:Z97052; PTDN:CA809759.1; GSPDB:GN00068; SPDB:SPCC4G3.14
A:Experimental source: strain 972h.; cosmid C4G3
A:Genetics:
A:Gene: SPDB:SPCC4G3.14
A:Map position: 3
A:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

F;86-150/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 38.4%; Score 48; DB 2; Length 528;
Best Local Similarity 52.6%; Pred. No. 24;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 3 GFRSARRRREPGLDTPK 21

Db 504 GFFKRAFRRLHPDEDQPK 522

RESULT 5

T38490
hypothetical protein SPAC29B12.02c - fission yeast (Schizosaccharomyces pombe)
C.Species: Schizosaccharomyces pombe
C.Date: 03-Dec-1999
C.Sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C.Accession: T38490
R.Gentles, S.; Chercher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
Submitted to the ENBL Data Library, September 1997
A.Reference number: Z21797

A1:Accession: T39490
 A1:Status: preliminary; translated from GB/EMBL/DBJ
 A1:Molecule type: DNA
 A1:Residues: 1-798 <GEN>
 A1:Cross-references: EMBL:Z99164; PIDN:CAB16247.1; GDB:
 A1:Experimental source: strain 972h; cosmid c29B12
 A1:Genetics:
 A1:Gens: SPDB:SPAC29B12.02c
 A1:Map position: 1

Query Match	38.4%;	Score 48;	DB 2;	Length 798;
Best Local Similarity	50.0%;	Pred. No. 37;		

Qy 9 RRRREGLDPTPKVLE 24

Db 355 RQRRKKGIDETSKIIIE 370

RESULT 6

I40485
 surfactin synthetase component I - *Bacillus subtilis*
 NR1:Alternate names: competence protein srfAA; surfactin production protein srfAA; surfactin C-species: *Bacillus subtilis*
 revision 12-Aug-1996
 C.Date: 12-Aug-1996
 #text change 03-Nov-2000
 I40485; B69718; S35517; A37323; S45967; A43705; S34985
 R.Comina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Venema, G.; van Sinderen, A.
 Mol. Microbiol. 8, 821-831, 1993
 Article: Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*
 Reference number: I40485; MUID:93360813; PMID:8355609

A;Accession: 140485
A;Status: translated from GB/EMBL/DDBT

A; Molecule type: DNA

A;Residues: 1-3588 <RES>
A:Cross-references: EMRL:

A; Experimental source: strain W168 derivative of JH642
References: EMP1:X70338; NID:G396480; FIDN:CAA49816.1; FID:G396481

Nature 390, 249-256, 1997

Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, S.; Kunitz, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, S.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.A.; Ehrlich, S.D.; Emerson, P.T.; Enian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Accession: B69716

Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

Reference number: A69580; PMID:98044033; PMID:9384377

A,Status: nucleic acid sequence not shown; translation not shown

A,Molecule type: DNA

A,Residues: 1-3588 <UN>

A,Cross-references: GB:Z910105; GB:AL009126; NID:g2632457; PIDN:CAB12142.1; PID:g2632634

A,Experimental source: strain 168

A,Experimental source: Fujishima, Y.; Corbell, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane, K.

A,Nucleic acid sequence not shown; translation not shown

A,Title: Nucleotide sequence of 5' portion of srfA that contains the region required for

A,Reference number: S35517; MUID:93181186; PMID:8441623

A,Accession: S35517

A,Status: nucleic acid sequence not shown; significant sequence differences

A,Molecule type: DNA

A,Cross-references: EMBL:D13262; NID:g216345; PID:g216346

A,Experimental source: strain 168 trpC2

A,Note: protein sequence not complete, the nucleotide sequence was submitted to the EMBI

A,Residues: 1-46, 'EV', 49-145, 'I', 147-150, 'L', 152, 'AN', 155-280, 'T', 282-307, 'SF', 310-384 <

A,Cross-references: GB:M59339; NID:g14366

A,Experimental source: R.R.Nakano, M.M.; Magnuson, R.; Myers, A.; Curry, J.; Grossman, A.D.; Zuber, P.

A,Residues: 173, 1770-1778, 1991

A,Title: srfA is an operon required for surfactin production, competence development, an

A,Reference number: A37323; MUID:91154134; PMID:1847909

A,Accession: A37323

A,Status: nucleic acid sequence not shown; not compared with conceptual translation

A,Molecule type: DNA

A,Residues: 1-46, 'EV', 49-145, 'I', 147-150, 'L', 152, 'AN', 155-280, 'T', 282-307, 'SF', 310-384 <

A,Cross-references: GB:M59339; NID:g14366

A,Experimental source: R.R.Nakano, M.M.; Xia, L.; Zuber, P.

A,Residues: 173, 5487-5493, 1991

A,Title: Transcription initiation region of the srfA operon, which is controlled by the

A,Reference number: A43705; MUID:91358326; PMID:1715856

A,Contents: annotation

A,Experimental source: submitted to the EMBL Data Library, March 1993

A,Reference number: S46967

A,Accession: S46967

A,Molecule type: DNA

A,Residues: 3249-3271, 'A', 3273-3316, 'R', 3318-3451, 'Y', 3453-3483, 'DE', 3486-3487, 'DAGL', 34

A,Cross-references: EMBL:X72672; NID:g516359; PIDN:CAA51222.1; PID:g516359

A,Experimental source: strain 168 trpC2

A,Comment: This enzyme is one of several in the multienzyme complex that synthesizes the

A,Genetics:

A,Gene: srfA; srfA1

A,Note: srfA is the first gene of the srfA operon

A,Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier protein h

A,Keywords: antibiotic biosynthesis; carrier protein; phosphopantetheine; phosphoprotein

A,F:507-956/Domain: acetate-CoA ligase homology <ACL1>

A,F:974-1042/Domain: acyl carrier protein homology <ACP1>

A,F:1043-1488/Domain: repeat <RPT1>

A,F:1549-1993/Domain: acetate-CoA ligase homology <ACL2>

A,F:2080-2527/Domain: repeat <RPT2>

A,F:2589-3025/Domain: acetate-CoA ligase homology <ACL3>

A,F:3042-3109/Domain: acyl carrier protein homology <ACP3>

A,F:1006,2043,3074/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 38.4%; Score 48; DB 2; Length 3588;

Best Local Similarity 45.5%; Pred. No. 1.7e+02;

Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 GFFRSARRRRREPGLDTPKYLE 24

DB 1093 GFRYSKAGARISGINTPRVIE 1114

RESULT 7

C83525

ColQ protein PA0969 [imported] - Pseudomonas aeruginosa (strain PAO1)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Nov-2001

C:Accession: C83525

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:AEOC4530; GB:AEOC4091; NID:g9946865; PIDN:AAG04358.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: tolQ; PA0969
C:Superfamily: biopolymer transport protein

Query Match 37.6%; Score 47; DB 2; Length 231;
Best Local Similarity 44.4%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDP 18
 : ||| : ||| : ||| :
Db 89 RAGFKFESRLQQPGVDP 106

RESULT 8
AI0961
conserved hypothetical protein STY3975 [imported] - *Salmonella enterica* subsp. *enterica*
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A>Note: This species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001.#text_change 18-Nov-2002
C:Accession: AI0961
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
et al. Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* sero-
var Typhimurium DT104. Molec Cell Proteomics 1:21-32, 2002
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AI0961
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03189.1; PID:g16504821; GSPDB:GN00176
C:Genetics:
A:Gene: STY3975

Query Match 36.8%; Score 46; DB 2; Length 120;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 RSARRRRREPGLDP 18
 ||| | ||| | ||| |
Db 5 RRARRNDPGLQP 17

RESULT 9
S78089
G-protein signaling regulator RGS3 - human
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 24-Nov-1999
C:Accession: S78089; S68436
R:Druey, K.
submitted to the EMBL Data Library, May 1995
A:Reference number: S78089
A:Accession: S78089
A:Molecule type: mRNA
A:Residues: 1-519 <DRU>
A:Cross-references: EMBL:U27655; NID:gl216368; PID:gl216369
R:Druey, K.M.; Blumberg, P.M.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A>Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene
product. Science 273:1503-1506, 1996
A:Reference number: A58012; MUID:96178495; PMID:8502223
A:Accession: S68436
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-355, 'K', 357-519 <DRW>

A:Cross-references: EMBL:U27655
 A:Experimental source: tonsil
 A>Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-1

Query Match 36.8%; Score 46; DB 2; Length 519;
 Best Local Similarity 52.4%; Pred. No. 47;
 Matches 11; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 1 KLGFFSARRRRRPGDPTPK 21
 |||||
 DB 355 KLGIFR--RRNESPCAPPACK 373

RESULT 10

167945
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) - rat (fragment)
 N:Alternate names: CAMP phosphodiesterase 1
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Nov-2000
 C:Accession: I67945; A33904
 R:Belger, G.B.; Rodgers, L.K.; Riggs, M.
 Gene 149, 237-244, 1994
 A>Title: Differential CNS expression of alternative mRNA isoforms of the mammalian genes
 A:Reference number: I53865; MUID:95047482; PMID:7958996
 A:Accession: I67945
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-536 <RES>
 A:Cross-references: GB:I27061; NID:9436011; PIDN:AAA56858.1; PID:9436012
 R:Swinnen, J.V.; Joseph, D.R.; Conti, M.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5325-5329, 1989
 A>Title: Molecular cloning of rat homologues of the Drosophila melanogaster dunce cAMP P
 A:Reference number: A33904; MUID:89315790; PMID:2546153
 A:Accession: A33904
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 230-496 <SWI>
 A:Cross-references: GB:M25347; GB:M25410
 C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',5'
 C:Keywords: alternative splicing; cAMP binding; phosphoric diester hydrolase
 F:253-481/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNPD>

Query Match 36.8%; Score 46; DB 2; Length 536;
 Best Local Similarity 52.9%; Pred. No. 49;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 FFRSARRRRRPGDPTPK 20
 |||||
 DB 434 FFGQGRERESGLDISP 450

RESULT 11

S42047
 aspartate-tRNA ligase (EC 6.1.1.12) - Mycobacterium leprae
 N:Alternate names: aspartyl-tRNA synthetase
 C:Species: Mycobacterium leprae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S42047
 R:Spierings, E.H.T.; Wiele, B.; Thole, J.E.R.
 Submitted to the EMBL data Library, February 1994
 A:Reference number: S42047
 A:Accession: S42047
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-589 <SPI>
 A:Cross-references: EMBL:X77655; NID:9454973; PIDN:CAA54735.1; PID:9581332
 C:Genetics:
 A:Gene: ASFP
 A:Start codon: GTG
 C:Function:
 A:Description: activates amino acid and transfers it to specific tRNA molecule
 A:Pathway: protein biosynthesis
 C:Superfamily: lysine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 Query Match 36.8%; Score 46; DB 1; Length 589;
 Best Local Similarity 44.4%; Pred. No. 54;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 7 SARRRRRPGDPTPKVLE 24
 |||||
 DB 572 TAQRKESGIDTKPEVE 589

RESULT 12

H95367
 probable histidine kinase, chemotaxis (EC 2.7.3.-) [imported] - Sinorhizobium meliloti (C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 22-Jun-2003
 C:Accession: H95367
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: H95367
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK65506.1; PID:gl4523979; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma1558
 A:Genome: plasmid
 C:Superfamily: chemotaxis hybrid-type signal transduction histidine kinase, PrZe type; r
 C:Keywords: phosphotransferase

Query Match 36.8%; Score 46; DB 2; Length 708;
 Best Local Similarity 60.0%; Pred. No. 65;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 ARRRRRRPGDPTPKV 22
 |||||
 DB 204 ARKMRAPGSDPTGV 218

RESULT 13

S71626
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), cAMP-specific - human
 N:Alternate names: 3',5'-cyclic AMP phosphodiesterase
 C:Species: Homo sapiens (man)
 C>Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 17-Nov-2000
 C:Accession: S71626; I61356
 R:Engels, P.; Sullivan, M.; Mueller, T.; Luebbert, H.
 FEBS Lett. 358, 305-310, 1995
 A>Title: Molecular cloning and functional expression in yeast of a human cAMP-specific p
 A:Reference number: S71626; MUID:95145731; PMID:7843419
 A:Accession: S71626
 A:Molecule type: DNA
 A:Residues: 1-712 <ENG>
 A:Cross-references: EMBL:Z46632; NID:g727222; PIDN:CAA86601.1; PID:g727223
 A:Experimental source: substantia nigra
 R:Belger, G.; Michaeli, T.; Martins, T.; St. John, T.; Steiner, B.; Rodgers, L.; Riggs,
 Mol. Cell. Biol. 13, 6588-6571, 1993
 A>Title: A family of human phosphodiesterases homologous to the dunce learning and memor
 A:Reference number: A54442; MUID:94019330; PMID:8413254
 A:Accession: I61356

A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 462-712 <RES>
A:Cross-references: GB:L20968; NID:g347125; PIDN:AAA03591.1; PID:g347126
C:Genetics:
A:Gene: HSPD4C1
C:Function:
A:Description: cAMP hydrolysis; converts cAMP to non-cyclic AMP
A:Pathway: cyclic nucleotide metabolism
A>Note: expressed in various tissues but not in cells of the immune system
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',5'
C:Keywords: phosphoric diester hydrolase
F:387-615/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 36.8%; Score 46; DB 2; Length 712;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 FFRSARRRRPGLDPTP 20
||: ||| ||| ||| :
Db 568 FFOQGDREERESGLDISP 584

RESULT 14
G84384
DNA polymerase type II large chain [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84384
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84384
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1370 <STO>
A:Cross-references: GB:AE004437; NID:g10581749; PIDN:AAG20443.1; GSPDB:GN00138
C:Genetics:
A:Gene: polA2

Query Match 36.8%; Score 46; DB 2; Length 1370;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 FRSARRRRPGLDPTPKV 22
| | | | | | | | | |
Db 20 FEVAEARRERGEDPTFNV 37

RESULT 15
B33904
cAMP phosphodiesterase 2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 02-Sep-2000
C:Accession: B33904
R;Swinnen, J.V.; Joseph, D.R.; Conti, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 5325-5329, 1989
A:Title: Molecular cloning of rat homologues of the Drosophila melanogaster dunce cAMP P
A:Reference number: A33904; MUID:89315790; PMID:2546153
A:Accession: B33904
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-267 <SWI>
A:Cross-references: GB:M25348
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase homology
F:24-252/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 36.0%; Score 45; DB 2; Length 267;
Best Local Similarity 47.1%; Pred. No. 34;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 4 FFRSARRRRPGLDPTP 20
||: ||| ||| ||| :
Db 205 FFOQAHRRERGRGWEISP 221

Search completed: September 21, 2004, 13:05:04
Job time : 5.2067 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:44:02; Search time 0.703911 Seconds
(without alignments)
1775.344 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1189
Perfect score: 125
Sequence: 1 KLGFFRSARRRRPEGLDTPKVL 24

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	1189	1 ITAH_HUMAN	Q9ukx5 homo sapien
2	52	41.6	956	1 GLK4 RAT	Q01812 rattus norv
3	49	39.2	411	1 DPO4 HALN1	Q9hqt4 halobacteri
4	48	38.4	3581	1 SRFI_BACSU	P27206 bacillus su
5	47	37.6	231	1 TOLA_PSEAE	P50598 pseudomonas
6	47	37.6	262	1 ORC6 MOUSE	Q9wuj8 mus musculu
7	46	36.8	519	1 RGS3 HUMAN	P49796 homo sapien
8	46	36.8	536	1 CN4C RAT	P14644 rattus norv
9	46	36.8	589	1 SYD_MYCLE	P36429 mycobacteri
10	46	36.8	712	1 CN4E_HUMAN	Q08493 homo sapien
11	46	36.8	1370	1 DP2L_HALN1	Q9hmx8 halobacteri
12	45	36.0	485	1 SOC6 MOUSE	Q8vhq2 mus musculu
13	45	36.0	579	1 SYFB_THETN	Q8r9c7 thermonaer
14	45	36.0	794	1 SYFB_THETN	P34433 caenorhabdi
15	44.5	35.6	244	1 QL53_CABEL	Q69583 mycobacteri
16	44	35.2	289	1 QRC8 MYCLE	Q91127 streptomyce
17	44	35.2	328	1 MPR2_STRFO	P43161 streptomyce
18	44	35.2	344	1 MPR2_STRFO	Q92866 rickettsia
19	44	35.2	554	1 Y478 RICPR	Q8twx0 methanopyru
20	44	35.2	564	1 CABI_METKA	Q14137 homo sapien
21	44	35.2	746	1 BOPI_HUMAN	Q8p7z6 xanthomonas
22	44	35.2	791	1 SYFB_XANCP	Q8p7e5 xanthomonas
23	44	35.2	792	1 SYFB_XANAC	Q8b7m5 mus musculu
24	44	35.2	825	1 NLG3_MOUSE	Q9nt94 homo sapien
25	44	35.2	848	1 NLG3_HUMAN	Q62889 rattus norv
26	44	35.2	848	1 NLG3_HUMAN	Q63679 rattus norv
27	44	35.2	1214	1 TSGA RAT	P23347 rattus norv
28	44	35.2	1234	1 B3A2 RAT	P13808 mus musculu
29	44	35.2	1237	1 B3A2_MOUSE	P48746 oryctolagus
30	44	35.2	1237	1 B3A2_RABIT	Q9z0s8 cavia porce
31	44	35.2	1238	1 B3A2_CAVPO	Q62261 mus musculu
32	44	35.2	2363	1 SPFO_MOUSE	P47810 mus musculu
33	43.5	34.8	646	1 WEE1_MOUSE	

RESULT 1

ITAH_HUMAN

ID ITAH_HUMAN STANDARD; PRT; 1189 AA.

AC Q9UKX5; Q9UKQ1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-11 precursor.

GN ITGA11.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal heart, and Osteoblast;

RX MEDLINE=99417679; PubMed=10486209;

RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,

RA Wang S.-X., Morris C.M., Kriesansen G.W.;

RT "Cloning, sequence analysis, and chromosomal localization of the novel

RT human integrin alpha11 subunit (ITGA11).";

RL Genomics 60:179-187(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal muscle, and Uterus;

RX MEDLINE=99395147; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;

RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)

RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated

RT integrin alpha-chain present in muscle tissues.";

RL J. Biol. Chem. 274:25735-25742(1999).

RN [3]

RP SEQUENCE OF 954-1188 FROM N.A.

RC TISSUE=Fibroblast;

RA Andreu N., Estivill X., Escarceller M., Sumoy L.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC -!- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11

CC ASSOCIATES WITH BETA-1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND

CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO

CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO

CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,

CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN

CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.

CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING

CC FETAL MUSCLE CELLS (IN VITRO).

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 7 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC

Q63802 rattus norv
Q60840 homo sapien
P14270 rattus norv
Q08499 homo sapien
Q14302 schizosacch
P71533 mycobacteri
P04920 homo sapien
Q9J187 mus musculu
Q88480 rattus norv
Q967v3 homo sapien
Q8pnt4 xanthomonas
Q8rtj5 xanthomonas

ALIGNMENTS


```

CC EMBL; X59996; CA442615.1; -
DR EMBL; U08257; AAA17830.1; -
DR PIR; J50685; J50685.
DR HSSP; P19491; IGR2.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR001320; Ion_glu_receptor.
DR InterPro; IPR001622; K-channel_pore.
DR InterPro; IPR001508; NMDA_receptor.
DR InterPro; IPR001311; SBP/Glu_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
DR Pfam; PF00660; lig_chan; 1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PBPE; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Signal; Transmembrane.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 956 GLUTAMATE RECEPTOR, IONOTROPIC KAINATE 4.
FT DOMAIN 21 545 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 546 566 POTENTIAL.
FT TRANSMEM 624 644 POTENTIAL.
FT TRANSMEM 805 825 POTENTIAL.
FT DOMAIN 862 865 POLY-ARG.
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 473 479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 539 539 S -> F (IN REF. 2).
SQ SEQUENCE 956 AA; 107223 MW; 970C9D72C8D74700 CRC64;

Query Match 41.6%; Score 52; DB 1; Length 956;
Best Local Similarity 68.8%; Pred. No. 5.6;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 RRRRFGGLDPTPKVL 24
||||| ||||| |||||
DB 862 RRRRSGGLPQPFVLE 877

RESULT 3
ID DPO4_HALN1 STANDARD; PRT; 411 AA.
AC Q9HQ74;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase IV (EC 2.7.7.7) (Pol IV).
DB DBH OR VNG1014G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2050448; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.B., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Poorly processive, error-prone DNA polymerase involved
CC in untargeted mutagenesis. Copies undamaged DNA at stalled
CC replication forks, which arise in vivo from mismatched or
CC misaligned primer ends. These misaligned primers can be extended
CC by polIV. Exhibits no 3'-5' exonuclease (proofreading) activity.

```

```

CC May be involved in translational synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.
CC -!- SIMILARITY: Contains 1 umuc domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005036; AAG19429.1; ALT_INIT.
CC HMAP; MF 01113; -. 1.
CC InterPro; IPR001126; UMUC_like.
CC Pfam; PF00817; IMS; 1.
CC PROSITE; PS00173; UMUC; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding; Mutator protein; Magnesium; Complete proteome.
FT DOMAIN 18 211 UMUC.
FT SITE 27 27 SUBSTRATE DISCRIMINATION (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
FT METAL 22 22 MAGNESIUM (BY SIMILARITY).
FT METAL 130 130 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 411 AA; 43492 MW; CA152C89EE8EAF81 CRC64;

Query Match 39.2%; Score 49; DB 1; Length 411;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 FFRSARRRREGGLDPTPKVL 23
||||| ||||| |||||
DB 26 FYASCERRREPALRGAPVV 45

RESULT 4
SRF1_BACSU STANDARD; PRT; 3588 AA.
ID SRF1_BACSU STANDARD; PRT; 3588 AA.
AC P27206;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Surfactin synthetase subunit 1.
GN SRF4A OR SRF4I OR SRF4 OR BSU03480.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93181186; PubMed=8441623;
RA Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,
RA Zuber P., Yamane K.;
RA "Nucleotide sequence of 5' portion of srfA that contains the region
RT required for competence establishment in Bacillus subtilis."
RL Nucleic Acids Res. 21:93-97(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=93360813; PubMed=8355609;
RA Cosmina P., Rodriguez F., de Perra F., Grandi G., Perego M.,
RA Venema G., van Sinderen D.;
RT "Sequence and analysis of the genetic locus responsible for surfactin
RT synthesis in Bacillus subtilis."
RL Mol. Microbiol. 8:821-831(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;

```

RX MEDLINE=97124189; PubMed=8969502;
 RA Yanane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis
 chromosome: determination of the sequence of a 146 kb segment and
 identification of 113 genes.";
 RL Microbiology 142:3047-3056(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.D., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haiteh J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel D., Kasahara Y., Kluwe S., Kluwe S., Kluwe S.,
 RA Joris B., Karamata D., Kasahara Y., Kluwe S., Kluwe S., Kluwe S.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapilus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker F.,
 RA Winters P., Wisat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 RN [5]
 RP SEQUENCE OF 1-460 FROM N.A.
 RX MEDLINE=91154134; PubMed=1847909;
 RA Nakano M.M., Magnusson R., Myers A.M., Curry J., Grossman A.D.,
 RA Zuber P.;
 RT "srfA is an operon required for surfactin production, competence
 development, and efficient sporulation in Bacillus subtilis.";
 RL J. Bacteriol. 173:1770-1778(1991).
 RN [6]
 RP SEQUENCE OF 1-64 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95219080; PubMed=7704255;
 RA Fujishima Y., Yamane K.;
 RT "A 10 kb nucleotide sequence at the 5' flanking region (32 degrees)
 of srfAA of the Bacillus subtilis chromosome.";
 RL Microbiology 141:277-279(1995).
 RN [7]
 RP SEQUENCE OF 1-38 FROM N.A.
 RX MEDLINE=91358326; PubMed=1715856;
 RA Nakano M.M., Xia L., Zuber P.;
 RT "Transcription initiation region of the srfA operon, which is
 controlled by the comp-cmA signal transduction system in Bacillus
 subtilis.";
 RL J. Bacteriol. 173:5487-5493(1991).
 CC -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO
 CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.
 CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.
 CC -!- COFACTOR: Contains 3 covalently bound phosphopantetheines.
 CC -!- PATHWAY: Cyclic peptide antibiotic surfactin biosynthesis.
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC -!- SIMILARITY: Contains 3 acyl carrier domains.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D13262; BAA02522.1; -
 CC EMBL; X70356; CAA49816.1; -
 CC EMBL; D50453; BAA08982.1; -
 CC EMBL; Z39105; CAB12142.1; -
 CC EMBL; M59939; AAA22815.1; -
 CC EMBL; D30762; BAA21034.1; -
 CC EMBL; M64702; AAA22816.1; -
 CC PIR; I40485; I40485.
 CC HSSP; P14687; I40485.
 CC Subtilisin; BGI010168; srfAA
 CC InterPro; IPR000873; AMP-bind.
 CC InterPro; IPR001242; Condensatn.
 CC InterPro; IPR006163; Pp_bind.
 CC InterPro; IPR006162; Ppantne S.
 CC Pfam; PF00501; AMP-binding; 3.
 CC Pfam; PF00668; Condensation; 4.
 CC Pfam; PF00550; pp-binding; 3.
 CC PRINTS; PR00154; AMPBINDING.
 CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 CC PROSITE; PS00455; AMP BINDING; 3.
 CC PROSITE; PS00075; ACP DOMAIN; 3.
 CC KW Ligase; Antibiotic biosynthesis; Phosphopantetheine; Sporulation;
 CC Multi-functional enzyme; Repeat; Complete proteome.
 CC FT REPEAT ? 1047 DOMAIN 1 (GLU-ACTIVATING).
 CC FT REPEAT ? 2084 DOMAIN 2 (LEU-ACTIVATING).
 CC FT REPEAT ? 3115 DOMAIN 3 (D-LEU-ACTIVATING).
 CC FT DOMAIN 976 1043 ACYL CARRIER (ACP) 1.
 CC FT DOMAIN 2013 2080 ACYL CARRIER (ACP) 2.
 CC FT DOMAIN 3044 3110 ACYL CARRIER (ACP) 3.
 CC FT BINDING 1006 1006 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT BINDING 2043 2043 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT BINDING 3074 3074 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT CONFLICT 146 146 M -> I (IN REF. 1 AND 5).
 CC FT CONFLICT 151 151 V -> L (IN REF. 1 AND 5).
 CC FT CONFLICT 165 165 G -> A (IN REF. 1).
 CC FT CONFLICT 281 281 Q -> T (IN REF. 1 AND 5).
 CC FT CONFLICT 460 460 D -> T (IN REF. 5).
 CC FT CONFLICT 540 540 A -> P (IN REF. 1).
 CC FT CONFLICT 562 562 Y -> I (IN REF. 1).
 CC FT CONFLICT 639 640 PT -> GS (IN REF. 1).
 CC FT CONFLICT 644 644 P -> R (IN REF. 1).
 CC FT CONFLICT 647 649 FLIP -> LLA (IN REF. 1).
 CC FT CONFLICT 1026 1026 Q -> L (IN REF. 1).
 CC FT CONFLICT 1065 1114 HWSHSHSSASSRWSSESAIHSSNSSEGFYSKAGARISG
 CC INPTRVIE -> SLAQRTYIVSQPDAGVGNPPAAAIL
 CC EGPLDQKLERAFQGLIRHESLR (IN REF. 1).
 CC VC -> DS (IN REF. 1).
 CC A -> R (IN REF. 1).
 CC L -> V (IN REF. 1).
 CC T -> S (IN REF. 1).
 CC T -> D (IN REF. 1).
 CC A -> R (IN REF. 1).
 CC PAVFQMD -> LRCLSKWT (IN REF. 1).
 CC P -> L (IN REF. 1).
 CC Q -> H (IN REF. 1).
 CC HR -> QQ (IN REF. 1).
 CC L -> V (IN REF. 1).
 CC C -> S (IN REF. 1).
 CC S -> R (IN REF. 1).
 CC AV -> RC (IN REF. 1).
 CC MISSING (IN REF. 1).
 CC S -> T (IN REF. 1).
 CC TA -> SP (IN REF. 1).
 CC N -> P (IN REF. 1).
 CC CONFLICT 1130 1131
 CC CONFLICT 1162 1162
 CC CONFLICT 1456 1456
 CC CONFLICT 1848 1848
 CC CONFLICT 1892 1892
 CC CONFLICT 1909 1909
 CC CONFLICT 1971 1978
 CC CONFLICT 2050 2050
 CC CONFLICT 2216 2216
 CC CONFLICT 2263 2264
 CC CONFLICT 2289 2289
 CC CONFLICT 2347 2347
 CC CONFLICT 2426 2426
 CC CONFLICT 2609 2610
 CC CONFLICT 2684 2686
 CC CONFLICT 2756 2756
 CC CONFLICT 2897 2898
 CC CONFLICT 3026 3026

```

FT CONFLICT 3097 3097 N -> F (IN REF. 1).
FT CONFLICT 3272 3272 S -> A (IN REF. 1).
FT CONFLICT 3317 3317 S -> R (IN REF. 1).
FT CONFLICT 3452 3452 S -> Y (IN REF. 1).
FT CONFLICT 3484 3484 HQMSHPPTTUSH -> DEMSDAGLITRSE (IN REF. 1).
FT CONFLICT 3501 3501 PH -> GO (IN REF. 1).
FT CONFLICT 3508 3508 E -> R (IN REF. 1).
FT CONFLICT 3588 AA; 403072 MW; BC02FBI57D7F1FDB CRC64;
SQ SEQUENCE 3588 AA; 403072 MW; BC02FBI57D7F1FDB CRC64;

Query Match 38.4%; Score 48; DB 1; Length 3588;
Best Local Similarity 45.5%; Pred. No. 92;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 GFPSARRRRRPGLDPTPKVLE 24
DB 1093 GFYSKAGARISGINTPRVIE 1114

RESULT 5
TOLQ_PSEAE STANDARD; PRT; 231 AA.
AC P50598;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TOLQ protein.
GN TOLQ OR PA0969.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAO;
RX MEDLINE=97113525; PubMed=8955385;
RA Dennis J.J., Lafontaine E.R., Sokol P.A.;
RT "Identification and characterization of the tolQRA genes of
Pseudomonas aeruginosa."
RL J. Bacteriol. 178:7059-7068 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen."
RL Nature 406:959-964 (2000).
CC -!- FUNCTION: Involved in the tonB-independent uptake of proteins (By
similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Potential).
CC -!- SIMILARITY: Belongs to the exbB / tolQ family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U39558; AAC44658.1; -.
CC EMBL; AE004530; BAG04359.1; -.
CC PIR; C83525; C83525.
CC InterPro; IPR002898; MotA_ExbB.
CC Pfam; PF01618; MotA_ExbB; 1.
CC Transport; Protein transport; Transmembrane; Inner membrane;
KW
```

```

KW Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
SQ SEQUENCE 231 AA; 25282 MW; 62E22A7E3B93C09D CRC64;

Query Match 37.6%; Score 47; DB 1; Length 231;
Best Local Similarity 44.4%; Pred. No. 7.1;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLGFPSARRRRRPGLDPP 18
DB 89 RAGFKFSRLRQFGVDP 106

RESULT 6
ORC6_MOUSE STANDARD; PRT; 262 AA.
ID ORC6_MOUSE
AC Q9WU08;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Origin recognition complex subunit 6.
GN ORC6L OR ORC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Dean F.B., O'Donnell M.;
RT "cDNA Cloning of a homolog for Saccharomyces cerevisiae ORC6 from Mus
musculus."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the origin recognition complex (ORC) that
binds origins of replication. It has a role in both chromosomal
replication and mating type transcriptional silencing. Binds to
the ASS consensus sequence (ACS) of origins of replication in an
ATP-dependent manner (By similarity).
CC -!- SUBUNIT: ORC is composed of six subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the ORC6 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF139659; AAD32667.1; -.
CC MGD; MGI:1329285; Orc6l.
CC InterPro; IPR008721; ORC6.
CC Pfam; PF05460; ORC6; 1.
CC DNA replication; Nuclear protein; DNA-binding.
KW SEQUENCE 262 AA; 29188 MW; F8D27BF9C87DE16C CRC64;

Query Match 37.6%; Score 47; DB 1; Length 262;
Best Local Similarity 40.0%; Pred. No. 8.1;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 5 FRSARRRRRPGLDPTPKVLE 24
DB 205 FSPTLKKFEGLEPPAKEIE 224

RESULT 7
RGS3_HUMAN STANDARD; PRT; 519 AA.
ID RGS3_HUMAN
AC P49756; Q8TD59; Q8TD68;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
KW
```

15-MAR-2004 (Rel. 43, Last annotation update)
Regulator of G-protein signaling 3 (RGS3) (RG3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=96178495; PubMed=8602223;
RA Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT "Inhibition of G-protein-mediated MAP Kinase activation by a new
mammalian gene family."
RL Nature 379:742-746(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RA Publ H.L. III, Ikeda S.R., Aronstam R.S.;
RT "cDNA clones of human proteins involved in signal transduction
sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
activity of G protein alpha subunits thereby driving them into
their inactive GDP-bound form.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P49796-1; Sequence=Displayed;
CC Name=2; Synonyms=RGS3f;
CC IsoId=P49796-2; Sequence=VSP_005662;
CC -!- PFM: Phosphorylated by cyclic GMP-dependent protein kinase (By
similarity).
CC -!- SIMILARITY: Contains 1 RGS domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U27655; AAC50394.1; -;
CC EMBL; AF493927; AAM12641.1; -;
CC EMBL; AF493941; AAM12655.1; -;
CC EMBL; AL162727; CAC78977.1; -;
CC PIR; S78089; S78089.
CC HSP; P49799; IAGR.
CC Genew; HGNC:9999; RGS3.
CC MIM; 602189; -;
CC GO; GO:0005829; Cytosol; TAS.
CC GO; GO:0005096; P:GTPase activator activity; TAS.
CC GO; GO:0001086; P:inactivation of MAPK; TAS.
CC GO; GO:0008277; P:regulation of G-protein coupled receptor pr. .; TAS.
CC InterPro; IPR000342; Regl_Grotein.
CC Pfam; PF00615; RGS; 1.
CC PRINTS; PR01301; RGS-PROTEIN.
CC ProDom; PD001580; Regl_Grotein; 1.
CC SMART; SM00315; RGS; 1.
CC PROSITE; PS00132; RGS; 1.
CC Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT DOMAIN 394 510 RGS.
FT VARSLIC 1 313 Missing (in isoform 2).
FT FTIC=VSP_005662.
FT CONFLICT 305 305 K -> R (IN REF. 2).
FT SEQUENCE 519 AA; 56601 MW; F1CFE3F2/D4673A0 CRC64;
Query Match 36.8%; Score 46; DB 1; Length 519;
Best Local Similarity 52.4%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
QY 1 KLGFRRSARRRRPGLDPTPK 21
DB 355 KLGIFR--RRNESPGAPGAK 373
RESULT 8
CNAC_RAT
ID CN4C_RAT STANDARD; PRT; 536 AA.
AC P14644;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CAMP-specific 3',5'-cyclic phosphodiesterase 4C (EC 3.1.4.17)
DE (DPDE1) (Fragment).
GN PDE4C.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95047482; PubMed=7958996;
RA Bolger G.B., Rodgers L., Riggs M.;
RT "Differential CNS expression of alternative mRNA isoforms of the
mammalian genes encoding cAMP-specific phosphodiesterases";
RL Gene 149:237-244(1994).
RN [2]
RP SEQUENCE OF 153-511 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=89315790; PubMed=2546153;
RA Swinnen J.V., Joseph D.R., Conti M.;
RT "Molecular cloning of rat homologues of the Drosophila melanogaster
dunce cAMP phosphodiesterase: evidence for a family of genes";
RL Proc Natl. Acad. Sci. U.S.A. 86:5325-5329(1989).
CC -!- CATALYTIC ACTIVITY: Adenosine 3',5'-cyclic phosphate + H(2)O =
adenosine 5'-phosphate.
CC -!- ENZYME REGULATION: Inhibited by rolipram.
CC -!- PATHWAY: Cyclic nucleotide metabolism.
CC -!- SIMILARITY: Belongs to the cyclic nucleotide phosphodiesterase
family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L27061; AAA56858.1; -;
CC EMBL; M25347; AAA41847.1; -;
CC PIR; I67945; I67945.
CC InterPro; IPR003607; Met_phosphohydro.
CC InterPro; IPR002073; PDEase.
CC Pfam; PF00233; PDEase; 1.
CC PRINTS; PR00387; EDIESTERASE1.
CC SMART; SM00471; Hdc; 1.
CC PROSITE; PS00126; PDEASE_I; 1.
CC Hydrolase; cAMP; Multigene family.
FT NON_TER 1 1
FT DOMAIN 524 534 POLY-GLU.
FT CONFLICT 218 218 R -> S (IN REF. 2).
FT CONFLICT 507 507 S -> N (IN REF. 2).
FT SEQUENCE 536 AA; 60063 MW; 87D12BE2C4642F3 CRC64;
Query Match 36.8%; Score 46; DB 1; Length 536;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 4 FRRSARRRRPGLDPTPK 20
DB 355 KLGIFR--RRNESPGAPGAK 373

434 FFQGDREERESGLDIP 50

DR TIGRFAMs; TIGR00459; asps_bact; 1.
DE PROSITE; PS50862; AA_TRNA_LIGASE II; 1.
KW Aminoacyl-tRNA synthetase, Protein biosynthesis; Ligase; ATP-binding;
SQ Complete proteome.

Query Match 36.8%; Score 46; DB 1; Length 589;
Best Local Similarity 44.4%; Pred. No. 27;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps
SEQUENCE 589 AA; 64630 MW; 530EA0675357031C CRC64;

QY 7 SARRRRPGLDPTPKVLE 24
DB 572 TAQRKESGIDTKPEVE 589

RESULT 10
CN4C_HUMAN
ID CN4C_HUMAN STANDARD; PRT; 712 AA.
AC Q08493; Q9UN44; Q9UN45; Q9UN46; Q9UPJ6;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CAMP-specific 3',5'-cyclic phosphodiesterase 4C (EC 3.1.4.17)
DE (DPB1) (PDB1).
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM PDE4C1).
RC TISSUE=Substantia nigra;
RX MEDLINE=95145731; PubMed=7843419;
RA Engels P., Sullivan M., Mueller T., Luebbert H.;
RT "Molecular cloning and functional expression in yeast of a human
RL CAMP-specific phosphodiesterase subtype (PDE IV-C).";
FEBS Lett. 358:305-310(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS PDE4C1; PDE4C2 AND PDE4C3).
RX MEDLINE=20039485; PubMed=10574328;
RA Sullivan M., Olsen A.S., Houslay M.D.;
RT "Genomic organisation of the human cyclic AMP-specific
RT phosphodiesterase PDE4C gene and its chromosomal localisation to
RT 19p13.1 between RAB3A and JUND".
RL Cell. Signal. 11:735-742(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerding J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Carnes J.,
RA Panagoulias L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C".
RL Submitted (Oct-1998) to the EMBL/GenBank/DBAJ databases.
RN [4]
RP SEQUENCE OF 462-712 FROM N.A.
RX MEDLINE=94019330; PubMed=8413254;
RA Bolger G., Michaeli T., Martins T., St John T., Steiner B.,
RA Rodgers L., Riggs M., Wigler M., Ferguson K.;
RT "A family of human phosphodiesterases homologous to the dunce
RT learning and memory gene product of Drosophila melanogaster are
RT potential targets for antidepressant drugs".
RL Mol. Cell. Biol. 13:6558-6571(1993).
CC -!- CATALYTIC ACTIVITY: Adenosine 3',5'-cyclic phosphate + H(2)O =
CC adenosine 5'-phosphate.
CC -!- ENZYME REGULATION: Inhibited by rolipram.
CC -!- PATHWAY: Cyclic nucleotide metabolism.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=7;

SYD_MYCLE
STANDARD; PRT; 589 AA.
P36429; P95671;
01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--trna ligase)
(ASPRS) (Antigen TS).
ASPS OR MLC0501 OR MLCB1259.19.
Myobacterium leprae.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
[1]
SEQUENCE FROM N.A.
MEDLINE=96071886; PubMed=7591123;
Wiesle B., Spierings E., van Noort J., Naafs B., Offringa R.,
Ottenhoff T.;
RT "Molecular characterization and T-cell-stimulatory capacity of
MYCOBACTERIUM LEPRAE antigen TS.";
Infect. Immun. 63:4682-4685(1995).
[2]
SEQUENCE FROM N.A.
STRAIN=TNI.
MEDLINE=21128732; PubMed=11234002;
Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R.M., Devlin K., Duthoy S., Fellwell T., Fraser A., Hamlin N.,
Hollroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
Nature 409:1007-1011(2001).
-!- CATALYTIC ACTIVITY: ATP + L-aspartate + trna(asp) = AMP +
aspartate + L-asparlyl-trna(asp).
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X77655; CA54735.1; -;
EMBL; S82268; AAC27132.1; -;
EMBL; AL023591; CAAL9094.1; -;
EMBL; AL583918; CAC30009.1; -;
PIR; S42047; S42047.
HSP; P36419; 1EFW.
Leptoma; ML0501; -;
HAMAP; NF_00044; -; 1.
InterPro; IPR004524; ASPS bact.
InterPro; IPR004115; GAD_dom.
InterPro; IPR008994; Nucleic acid OB.
InterPro; IPR004364; tRNA-synt_2.
InterPro; IPR002312; tRNA_synt_asp.
InterPro; IPR004365; tRNA_anti_
InterPro; IPR006195; tRNA_ligase_II.
Pfam; PF02938; GAD; 1.
Pfam; PF00152; tRNA-synt_2; 2.
Pfam; PF01336; tRNA_anti; 1.
PRINTS; PR01042; TRNASYNTHASP.

CC Name=PDE4C1;
 CC IsoId=Q08493-1; Sequence=Displayed;
 CC Name=PDE4C2;
 CC IsoId=Q08493-2; Sequence=VSP_004575;
 CC Name=PDE4C3;
 CC IsoId=Q08493-3; Sequence=VSP_004574;
 CC Name=PDE4C4;
 CC IsoId=Q08493-4; Sequence=Not described;
 CC Name=PDE4C5;
 CC IsoId=Q08493-5; Sequence=Not described;
 CC Name=PDE4C6;
 CC IsoId=Q08493-6; Sequence=Not described;
 CC Name=PDE4C7;
 CC IsoId=Q08493-7; Sequence=Not described;
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN VARIOUS TISSUES BUT NOT IN CELLS
 CC OF THE IMMUNE SYSTEM.
 CC -!- SIMILARITY: Belongs to the cyclic nucleotide phosphodiesterase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z46632; CAA86601.1; -;
 CC EMBL; AF157816; AAD47053.1; -;
 CC EMBL; AF157811; AAD47053.1; JOINED.
 CC EMBL; AF157814; AAD47053.1; JOINED.
 CC EMBL; AF157815; AAD47053.1; JOINED.
 CC EMBL; AF157816; AAD47054.1; -;
 CC EMBL; AF157812; AAD47054.1; JOINED.
 CC EMBL; AF157814; AAD47054.1; JOINED.
 CC EMBL; AF157815; AAD47054.1; JOINED.
 CC EMBL; AF157816; AAD47055.1; -;
 CC EMBL; AF157814; AAD47055.1; JOINED.
 CC EMBL; AF157815; AAD47055.1; JOINED.
 CC EMBL; AC005759; AAC83047.1; -;
 CC EMBL; L20968; AAA03591.1; -;
 CC PIR; S71626; S71626.
 CC PDB; 1LXU; 26-JUN-02.
 CC Genew; HGNC:18782; PDE4C.
 CC MIM; 600128; -;
 CC GO; GO:0004115; F:cAMP-specific phosphodiesterase activity; TAS.
 CC InterPro; IPR003607; Met phosphohydro.
 CC InterPro; IPR002073; PDEase.
 CC Pfam; PF00233; PDEase; 1.
 CC PRINTS; PR00387; PDIESTERASE1.
 CC SMART; SM00471; Hdc; 1.
 CC PROSITE; PS00126; PDEASE_I; 1.
 CC Hydroxylase; CAMP; Multigene family; Alternative splicing; 3D-structure.
 CC VARSPLIC 1 81
 CC MENLGVSDKAGACSRSLRSGHSMTRAPKHLWRQPREPIR
 CC IQQFVSDPKSAGCERDLSPRPFLKSLRSLMPSVSCRR
 CC -> MQGPAPAPVFGSPRSGSPRSGSPGLFRKLVLNQSLRL
 CC QRFVVARPLC (in isoform PDE4C3).
 CC /FTIG-VSP 004574.
 CC Missing (in isoform PDE4C2).
 CC VARSPLIC 1 106
 CC /FTIG-VSP 004575.
 CC K -> N (IN REF. 2).
 CC D -> Y (IN REF. 2).
 CC EL -> DV (IN REF. 1).
 CC NSE -> K (IN REF. 3).
 CC EL -> DV (IN REF. 1).
 CC SEQUENCE 712 AA; 79901 MW; 1932116C9CE0322C CRC64;
 CC -----
 CC Query Match 36.8%; Score 46; DB 1; Length 712;
 CC Best Local Similarity 52.9%; Pred. No. 33;
 CC Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 CC 4 PFRSARRRREPGLDTP 20
 CC ||: ||| ||| :|

DB 568 FFOQDRERESGLDISP 584
 RESULT 11
 ID DP2L_HALN1 STANDARD; PRT; 1370 AA.
 AC Q9HMX8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA polymerase II large subunit (EC 2.7.7.7) (Pol II) [Contains: Hsp-
 DE NRC1 polC intein (Hsp-NRC1 pol2 intein)].
 GN POLC OR POLA2 OR VNG2338G.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Dansen M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam W., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dasarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -!- FUNCTION: Possesses two activities: a DNA synthesis (polymerase)
 CC and an exonucleolytic activity that degrades single stranded DNA
 CC in the 3' to 5' direction. Has a template-primer preference which
 CC is characteristic of a replicative DNA polymerase (by similarity).
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA}(N).
 CC -!- CATALYTIC ACTIVITY: Degradation of single-stranded DNA. It acts
 CC progressively in a 3'- to 5'-direction, releasing nucleoside 5'-
 CC phosphates.
 CC -!- SUBUNIT: Heterodimer of a large subunit and a small subunit (By
 CC similarity).
 CC -!- PTM: This protein undergoes a protein self splicing that involves
 CC a post-translational excision of the intervening region (intein)
 CC followed by peptide ligation (potential).
 CC -!- SIMILARITY: Belongs to the archaeal DNA polymerase II family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB005116; AAC20443.1; -;
 CC PIR; G84384; G84384.
 CC HAMAP; MF_00324; -; 1.
 CC InterPro; IPR003587; Hedgehog hint N.
 CC InterPro; IPR003586; Hedgehog hint C.
 CC InterPro; IPR006141; Intein S.
 CC InterPro; IPR004475; PolC DF2.
 CC Pfam; PF03833; PolC_DP2; 1.
 CC SMART; SM00305; HintC; 1.
 CC SMART; SM00306; HintN; 1.
 CC TIGRFAMs; TIGR01443; Intein Cterm; 1.
 CC TIGRFAMs; TIGR01445; Intein Nterm; 1.
 CC TIGRFAMs; TIGR00354; PolC; 1.
 CC PROSITE; PS50818; INTEIN_C_TER; FALSE_NEG.
 CC PROSITE; PS50817; INTEIN_N_TER; 1.
 CC Transferrase; DNA-directed DNA polymerase; DNA replication; Hydroxylase;
 CC Nuclease; Exonuclease; DNA-binding; Multifunctional enzyme;
 CC Autocatalytic cleavage; Protein splicing; Complete proteome.
 CC CHAIN 1 925 DNA POLYMERASE II LARGE SUBUNIT, 1ST PART
 FT

```
FT CHAIN 926 1120 (POTENTIAL).
FT HSP-NRC1 POLC INTIN (POTENTIAL).
FT CHAIN 1121 1370 DNA POLYMERASE II LARGE SUBUNIT, 2ND PART
FT (POTENTIAL).
SQ SEQUENCE 1370 AA; 150295 MW; 07878AA9976790C9 CRC64;

Query Match
Best Local Similarity 36.8%; Score 46; DB 1; Length 1370;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 FRSARRRPPGLDPTPKV 22
DB 20 FEVAARERGEDPTFNV 37

RESULT 12
SOC6_HUMAN
ID SOC6_HUMAN STANDARD; PRT; 485 AA.
AC O1452;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Suppressor of cytokine signaling 6 (SOCS-6) (Nck, Ash and
DE phospholipase C gamma-binding protein) (Nck-associated protein 4)
DE (NAP-4) (Fragment).
DE SOCS6 OR NAP4.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98008966; PubMed=9344857;
RX "A novel ligand for an SH3 domain of the adaptor protein Nck bears an
RT SH2 domain and nuclear signaling motifs."
RL Biochem. Biophys. Res. Commun. 239:488-492(1997).
CC -!- FUNCTION: SOCS family proteins form part of a classical negative
CC feedback system that regulates cytokine signal transduction.
CC -!- SUBUNIT: Interacts, via the third proline-rich region, with the
CC second SH3 domain of the adaptor protein Nck. Also interacts with
CC Grb2 and phospholipase C-gamma.
CC -!- TISSUE SPECIFICITY: Expressed in brain and leukocytes. Also in
CC fetal lung fibroblasts and fetal brain.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SOCS box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005216; BAA22432.1; .
CC PIR; PC4427; PC4427.
CC HSSP; P23727; 2PNB.
CC GO; GO:0017124; F:SH3-domain binding; NAS.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001496; SOCS_C.
CC Pfam; PF00017; SH2; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00253; SOCS; 1.
CC PROSITE; PS50001; SH2; 1.
CC PROSITE; PS50225; SOCS; 1.
CC SH2 domain; Growth regulation; Signal transduction inhibitor.
CC
CC KW SH2 domain; Growth regulation; Signal transduction inhibitor.
CC
CC FT DOMAIN 82 402
CC PRO-RICH.
CC FT DOMAIN 301 381
CC SH2.
CC FT DOMAIN 398 507
CC SOCS BOX.
CC FT DOMAIN 502 552
CC POLY-PRO.
CC FT DOMAIN 84 97
CC POLY-GLY.
CC FT DOMAIN 141 149
CC POLY-GLN.
CC FT DOMAIN 181 185
CC POLY-PRO.
CC FT DOMAIN 186 195
CC POLY-PRO.
CC FT DOMAIN 301 310
CC POLY-PRO.
CC FT DOMAIN 341 348
CC POLY-PRO.
CC SQ SEQUENCE 579 AA; 62783 MW; FAB66BF2A0BE685A CRC64;

Query Match
Best Local Similarity 36.0%; Score 45; DB 1; Length 579;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRPPGLDPTP 20
```

```
FT DOMAIN 381 431 SOCS BOX.
FT DOMAIN 18 25 POLY-GLY.
FT DOMAIN 57 64 POLY-GLN.
SQ SEQUENCE 485 AA; 53564 MW; 0CC5EC107174A4F9 CRC64;

Query Match
Best Local Similarity 36.0%; Score 45; DB 1; Length 485;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRPPGLDPTP 20
DB 360 KNGKFLYFLRSRVFGLPPTP 379

RESULT 13
SOC6_MOUSE
ID SOC6_MOUSE STANDARD; PRT; 579 AA.
AC Q8VHQ2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Suppressor of cytokine signaling 6 (SOCS-6).
DE SOCS6 OR CISH7.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX Hilton D.J., Viney E.M., Alexander W.S., Willson T.A., Nicola N.A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SOCS family proteins form part of a classical negative
CC feedback system that regulates cytokine signal transduction.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SOCS box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF424814; AAL60516.1; .
CC MGD; MGI:1354911; Cish7.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001496; SOCS_C.
CC Pfam; PF00017; SH2; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00253; SOCS; 1.
CC PROSITE; PS50001; SH2; 1.
CC PROSITE; PS50225; SOCS; 1.
CC SH2 domain; Growth regulation; Signal transduction inhibitor.
CC
CC KW SH2 domain; Growth regulation; Signal transduction inhibitor.
CC
CC FT DOMAIN 82 402
CC PRO-RICH.
CC FT DOMAIN 301 381
CC SH2.
CC FT DOMAIN 398 507
CC SOCS BOX.
CC FT DOMAIN 502 552
CC POLY-PRO.
CC FT DOMAIN 84 97
CC POLY-GLY.
CC FT DOMAIN 141 149
CC POLY-GLN.
CC FT DOMAIN 181 185
CC POLY-PRO.
CC FT DOMAIN 186 195
CC POLY-PRO.
CC FT DOMAIN 301 310
CC POLY-PRO.
CC FT DOMAIN 341 348
CC POLY-PRO.
CC SQ SEQUENCE 579 AA; 62783 MW; FAB66BF2A0BE685A CRC64;

Query Match
Best Local Similarity 36.0%; Score 45; DB 1; Length 579;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRPPGLDPTP 20
```

```

DB      481 KNGFLYFLRGRVCLPPTP 500
SYFB_THETN
ID_ SYFB_THETN  STANDARD;      PRT;  794 AA.
AC  Q8R9C7;
DT  10-OCT-2003 (Rel. 42, Created)
DT  10-OCT-2003 (Rel. 42, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE  (Phenylalanyl-tRNA ligase beta chain) (PheRS).
GN  PHE1 OR TPE1868.
OS  Thermoanaerobacter tengcongensis.
OC  Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC  Thermoanaerobacteraceae; Thermoanaerobacter.
OX  NCBI_TaxID=119072;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MB4 / JCM 11007;
RX  MEDLINE=21992816; PubMed=1197336;
RA  Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA  Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA  Tan H., Chen R., Wang J., Yu J., Yang H.;
RT  "A complete sequence of T. tengcongensis genome.";
RL  Genome Res. 12:689-700(2002).
CC  -!- CATALYTIC ACTIVITY: AMP + L-phenylalanine + tRNA(Phe) = AMP +
CC  diphosphate + L-phenylalanyl-tRNA(Phe).
CC  -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC  -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC  similarity).
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
CC  family. Subfamily 1.
CC  -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC  -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AE013123; AAM24889.1; -
CC  HAMAP; MF_00283; -; 1
CC  InterPro; IPR005146; B3_4.
CC  InterPro; IPR005147; B5.
CC  InterPro; IPR005121; Fdx-AntiCB.
CC  InterPro; IPR008994; Nucleic acid_OB.
CC  InterPro; IPR004532; Phe1 bact.
CC  InterPro; IPR002547; tRNA_bind.
CC  Pfam; PF03483; B3_4; 1.
CC  Pfam; PF03484; B5; 1.
CC  Pfam; PF03147; FDX-ACB; 1.
CC  Pfam; PF01588; tRNA_bind; 1.
CC  TIGRFAMs; TIGR00472; phe1 bact; 1.
CC  PROSITE; P85086; TRBD; 1.
KW  Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW  Metal-binding; Magnesium; RNA-binding; tRNA-binding;
KW  Complete Proteome.
FT  DOMAIN 39 152 TRNA-BINDING.
FT  METAL 457 457 MAGNESIUM (BY SIMILARITY).
FT  METAL 463 463 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
FT  METAL 463 463 SIMILARITY).
FT  METAL 466 466 MAGNESIUM (BY SIMILARITY).
FT  METAL 467 467 MAGNESIUM (BY SIMILARITY).
SQ  SEQUENCE 794 AA; 88954 MW; F9179929B0461A7F CRC64;

Query Match 36.0%; Score 45; DB 1; Length 794;
Best Local Similarity 52.2%; Pred. No. 53;
Matches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

DB      1 KLGFFSARRRRPCLDPTPKVL 23
DB      353 KGLRSEASARFEKGLDPEITVL 375
RESULT 15
YL53 CAEEL
ID  YL53 CAEEL  STANDARD;      PRT;  244 AA.
AC  P34433;
DT  01-FEB-1994 (Rel. 28, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Hypothetical protein F44E2.3 in chromosome III.
DE  F44E2.3.
GN  Caenorhabditis elegans.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Bristol N2;
RX  MEDLINE=94150718; PubMed=7906398;
RA  Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA  Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
RA  Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA  Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA  Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA  Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA  Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA  Sims M., Smalton N., Smith A., Smith K., Sonhammer E., Staden K.,
RA  Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA  Wohldman P.;
RT  "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT  elegans.";
RL  Nature 368:32-38(1994).
CC  -!- SIMILARITY: TO DNAB.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; L23646; AAA28041.1; -
CC  PIR; S44822; S44822.
CC  WormPep; F44E2.3; CE00181.
CC  Hypothetical protein.
FT  DOMAIN 3 45 ARG/ASP/LYS-RICH.
FT  DOMAIN 79 90 PRO-RICH.
SQ  SEQUENCE 244 AA; 28994 MW; C0CB677FB01A2B18 CRC64;

Query Match 35.6%; Score 44.5; DB 1; Length 244;
Best Local Similarity 52.0%; Pred. No. 18;
Matches 13; Conservative 1; Mismatches 6; Indels 5; Gaps 2;

DB      1 KLG-FRRSARRRRPCLDPTTP 20
DB      58 QLGSIFFRRRRRERNESPRLPPPP 82
Search completed: September 21, 2004, 13:00:00
Job time : 2.70391 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:51:12 ; Search time 3.62011 Seconds
(without alignments)
2091.769 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFRRSARRRRPGLDTPKVL 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriapi:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	823	Q8WY18	Q8WY18 homo sapien
2	93	74.4	823	Q8CEB4	Q8CEB4 mus musculus
3	93	74.4	1188	Q7TQC3	Q7TQC3 mus musculus
4	54	43.2	958	Q8BWF5	Q8BWF5 mus musculus
5	50.5	40.4	59	Q804T8	Q804T8 fugu rubrip
6	50	40.0	179	Q87ZC4	Q87ZC4 pseudomonas
7	49	39.2	189	Q96N04	Q96N04 homo sapien
8	49	39.2	216	Q8VPO5	Q8VPO5 micrococccus
9	49	39.2	241	Q7WTF7	Q7WTF7 streptomyce
10	49	39.2	309	Q9GRD8	Q9GRD8 drosophila
11	49	39.2	472	Q9RX81	Q9RX81 streptomyce
12	49	39.2	740	Q8IZ41	Q8IZ41 homo sapien
13	49	39.2	947	Q8SYA1	Q8SYA1 drosophila
14	49	39.2	947	Q8MLT8	Q8MLT8 drosophila
15	49	39.2	1218	Q9V8R6	Q9V8R6 drosophila
16	48	38.4	168	Q7SVI3	Q7SVI3 gallus gall

17	48	38.4	283	13	Q7SVI2	Q7SVI2 gallus gall
18	48	38.4	373	16	O66737	O66737 aquifex aeo
19	48	38.4	408	13	Q8YI11	Q8YI11 gallus gall
20	48	38.4	441	13	Q7SVI0	Q7SVI0 gallus gall
21	48	38.4	528	3	P87239	P87239 schizosacch
22	48	38.4	798	3	O14026	O14026 schizosacch
23	48	38.4	799	13	Q7SVH9	Q7SVH9 gallus gall
24	47	37.6	66	16	O7UPB4	O7UPB4 rhodopirell
25	47	37.6	204	10	O94CP9	O94CP9 oryza sativ
26	47	37.6	231	2	O9WXX3	O9WXX3 pseudomonas
27	47	37.6	231	16	Q88N18	Q88N18 pseudomonas
28	47	37.6	237	4	Q8N0T0	Q8N0T0 homo sapien
29	47	37.6	248	10	Q8LN93	Q8LN93 oryza sativ
30	47	37.6	300	16	Q8XUM2	Q8XUM2 raistonia s
31	47	37.6	2144	2	Q8G982	Q8G982 planktothri
32	46.5	37.2	403	16	Q82AM6	Q82AM6 streptomyce
33	46	36.8	120	16	Q8ZL06	Q8ZL06 salmonella
34	46	36.8	120	16	Q8ZL06	Q8ZL06 salmonella
35	46	36.8	132	4	Q8WV02	Q8WV02 homo sapien
36	46	36.8	284	4	Q96NV5	Q96NV5 homo sapien
37	46	36.8	319	4	Q8NFN6	Q8NFN6 homo sapien
38	46	36.8	426	4	O9UPJ5	O9UPJ5 homo sapien
39	46	36.8	427	4	O76104	O76104 homo sapien
40	46	36.8	427	4	O43851	O43851 homo sapien
41	46	36.8	518	4	O43850	O43850 homo sapien
42	46	36.8	700	4	P78505	P78505 homo sapien
43	46	36.8	708	16	Q92YMS	Q92YMS rhizobium m
44	46	36.8	782	4	O76105	O76105 homo sapien
45	46	36.8	791	4	O43849	O43849 homo sapien

ALIGNMENTS

RESULT 1

Q8WY18	PRELIMINARY;	PRT;	823 AA.
ID	Q8WY18		
AC	Q8WY18;		
DT	01-MAR-2002 (TRENBLrel. 20, Created)		
DT	01-MAR-2002 (TRENBLrel. 20, Last sequence update)		
DE	01-JUN-2003 (TRENBLrel. 24, Last annotation update)		
DE	MSTP018.		
GN	MSTC18.		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI TaxID=9606;		
[1]			
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Adcta;		
RA	Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,		
RA	Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,		
RA	Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.F.,		
RA	Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;		
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF111799; AAL39001.1;		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	PRINTS; PR01185; INTEGRIN.		
DR	SMART; SM00191; Int. alpha; 4.		
SQ	SEQUENCE 823 AA; 92672 NW; DB4E78079DCD4925 CRC64;		

Query Match 100.0%; Score 125; DB 4; Length 823;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRPGLDTPKVL 24

DB 800 KLGFRRSARRRRPGLDTPKVL 823

```

RESULT 2
Q8CE84      PRELIMINARY;      PRT;      823 AA.
ID AC Q8CE84;
AC Q8CE84;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE MSTP018 homolog.
GN 4732459H24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK028821; BAC26137.1; -.
DR MGD; MGI:2442114; 4732459H24RIK.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
SQ SEQUENCE 823 AA; 92264 MW; A330236324A0E089 CRC64;

Query Match 74.4%; Score 93; DB 11; Length 823;
Best Local Similarity 70.8%; Pred. No. 1e-05;
Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRPPGLDPTPKVLE 24
DB 800 KLGFFRSARRRPPGLGPIPKELX 823

RESULT 3
Q7TQC3      PRELIMINARY;      PRT;      1188 AA.
ID AC Q7TQC3;
AC Q7TQC3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE All integrin.
GN ITGAll.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,
RA Guilberg D.;
RT "allbl integrin is important for mesenchymal cell function:
RT elimination of allbl leads to dwarfism.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Johansson M., Popova S.N.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV124460; ANM62130.1; -.
KW Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 74.4%; Score 93; DB 11; Length 1188;
Best Local Similarity 70.8%; Pred. No. 1.5e-05;

```

```

Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRPPGLDPTPKVLE 24
DB 1165 KLGFFRSARRRPPGLGPIPKELX 1188

RESULT 4
Q8BMF5      PRELIMINARY;      PRT;      956 AA.
ID AC Q8BMF5;
AC Q8BMF5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Glutamate receptor.
GN GRIK4 OR 6330551X01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK032029; BAC27660.1; -.
DR MGD; MGI:95817; Grik4.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005234; F:glutamate-gated ion channel activity; IEA.
DR GO; GO:0004970; F:ionotropic glutamate receptor activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005267; F:potassium channel activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR01828; ANF_receptor.
DR InterPro; IPR001320; Ion_glu_receptor.
DR InterPro; IPR001622; K_channel_pore.
DR InterPro; IPR001508; NMDA_receptor.
DR InterPro; IPR001311; SBP/Glu_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
DR Pfam; PF00060; Lig_chan; 1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PBPE; 1.
SQ SEQUENCE 956 AA; 107299 MW; 4A1E3951BFBA7A38 CRC64;

Query Match 43.2%; Score 54; DB 11; Length 956;
Best Local Similarity 68.8%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 RRRRPPGLDPTPKVLE 24
DB 862 RRRRSGGLPPQPVLE 877

RESULT 5
Q804T8      PRELIMINARY;      PRT;      59 AA.
ID AC Q804T8;
AC Q804T8;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Bone Gla protein.
GN BGP.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;

```

```

OC Tetradontoidae; Tetradontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Laize V., Canela M.L.;
RT "Identification of Takifugu rubripes osteocalcin (BGP) by comparative
   genomics.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF478914; AAC24898.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR002384; GLA_bone.
DR InterPro; IPR000294; VitA_gdp_GLA.
DR Pfam; PF00594; Gla; 1.
DR PRINTS; PRO0002; GLABONE.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
SQ SEQUENCE 59 AA; 6506 MW; 97112FFDA6A6AE1DB CRC64;

Query Match 40.4%; Score 50.5; DB 13; Length 59;
Best Local Similarity 68.8%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 RRRRREPGLDPTPKVLE 24
   ||||| |||: |||: |||
Db 11 RRRRAPG-EPTPQGLE 25

RESULT 6
ID Q87ZC4 PRELIMINARY; PRT; 179 AA.
AC Q87ZC4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome b561, putative.
GN "PSPTQ3505.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Barry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collmer A.;
RL "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016868; AAC056980.1; -.
DR TIGR; PSPTQ3505; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000516; Ni_hydr_CyTB.
DR Pfam; PF01292; Ni_hydr_CyTB; 1.
KW Complete proteome.
SQ SEQUENCE 179 AA; 20500 MW; 0F915D51AF24F92F CRC64;

Query Match 40.0%; Score 50; DB 16; Length 179;
Best Local Similarity 57.9%; Pred. No. 9.8;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 2 LGFRSARRRRR--PGLDPP 18
   ||||| |||: |||: |||
Db 61 LAFFRVAERRRQRPFGITP 79

RESULT 7
ID Q96N04 PRELIMINARY; PRT; 189 AA.
AC Q96N04;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ31614.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamiyama K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Osuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Nishio Y., Nagai K., Isogai T.;
RL "NEDO human cDNA sequencing project.";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056176; BAB71112.1; -.
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 2.
DR SMART; SM00854; EFh; 2.
DR PROSITE; PS00018; EF_HAND; 2.
KW Hypothetical protein.
SQ SEQUENCE 189 AA; 20736 MW; 472825570664F33D CRC64;

Query Match 39.2%; Score 49; DB 4; Length 189;
Best Local Similarity 52.2%; Pred. No. 15;
Matches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 LGFRSARRRRREPGLDPTPKVLE 24
   ||||| |||: |||: |||
Db 72 LGSIRGGRRDVGPLDPAPVSE 94

RESULT 8
ID Q8VPQ5 PRELIMINARY; PRT; 216 AA.
AC Q8VPQ5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MC16.
OS Micrococcus sp. 28.
OG Plasmid pSD10.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=161213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=28;
RA Zhong Z., Caspi R., Mincer T., Helinski D., Knauf V., Boardman K.,
RA Wilkinson J.E., Shea T., DeLoughery C., Toukdarian A.;
RT "A 50 kb plasmid rich in mobile gene sequences isolated from a marine
   Micrococcus.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034092; AAK62490.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 216 AA; 22533 MW; 7A082298C81297AF CRC64;

Query Match 39.2%; Score 49; DB 2; Length 216;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 6 RSARRRRREPGLDPTPKVLE 24
   ||||| |||: |||: |||
Db 165 RFSRRRRREPGLDPTPKVLE 183

```

1. **Introduction**
 2. **Background**
 3. **Methodology**
 4. **Results**
 5. **Discussion**
 6. **Conclusion**
 7. **References**
 8. **Appendix**
 9. **Figure 1**
 10. **Figure 2**
 11. **Figure 3**
 12. **Figure 4**
 13. **Figure 5**
 14. **Figure 6**
 15. **Figure 7**
 16. **Figure 8**
 17. **Figure 9**
 18. **Figure 10**
 19. **Figure 11**
 20. **Figure 12**
 21. **Figure 13**
 22. **Figure 14**
 23. **Figure 15**
 24. **Figure 16**
 25. **Figure 17**
 26. **Figure 18**
 27. **Figure 19**
 28. **Figure 20**
 29. **Figure 21**
 30. **Figure 22**
 31. **Figure 23**
 32. **Figure 24**
 33. **Figure 25**
 34. **Figure 26**
 35. **Figure 27**
 36. **Figure 28**
 37. **Figure 29**
 38. **Figure 30**
 39. **Figure 31**
 40. **Figure 32**
 41. **Figure 33**
 42. **Figure 34**
 43. **Figure 35**
 44. **Figure 36**
 45. **Figure 37**
 46. **Figure 38**
 47. **Figure 39**
 48. **Figure 40**
 49. **Figure 41**
 50. **Figure 42**
 51. **Figure 43**
 52. **Figure 44**
 53. **Figure 45**
 54. **Figure 46**
 55. **Figure 47**
 56. **Figure 48**
 57. **Figure 49**
 58. **Figure 50**
 59. **Figure 51**
 60. **Figure 52**
 61. **Figure 53**
 62. **Figure 54**
 63. **Figure 55**
 64. **Figure 56**
 65. **Figure 57**
 66. **Figure 58**
 67. **Figure 59**
 68. **Figure 60**
 69. **Figure 61**
 70. **Figure 62**
 71. **Figure 63**
 72. **Figure 64**
 73. **Figure 65**
 74. **Figure 66**
 75. **Figure 67**
 76. **Figure 68**
 77. **Figure 69**
 78. **Figure 70**
 79. **Figure 71**
 80. **Figure 72**
 81. **Figure 73**
 82. **Figure 74**
 83. **Figure 75**
 84. **Figure 76**
 85. **Figure 77**
 86. **Figure 78**
 87. **Figure 79**
 88. **Figure 80**
 89. **Figure 81**
 90. **Figure 82**
 91. **Figure 83**
 92. **Figure 84**
 93. **Figure 85**
 94. **Figure 86**
 95. **Figure 87**
 96. **Figure 88**
 97. **Figure 89**
 98. **Figure 90**
 99. **Figure 91**
 100. **Figure 92**
 101. **Figure 93**
 102. **Figure 94**
 103. **Figure 95**
 104. **Figure 96**
 105. **Figure 97**
 106. **Figure 98**
 107. **Figure 99**
 108. **Figure 100**
 109. **Figure 101**
 110. **Figure 102**
 111. **Figure 103**
 112. **Figure 104**
 113. **Figure 105**
 114. **Figure 106**
 115. **Figure 107**
 116. **Figure 108**
 117. **Figure 109**
 118. **Figure 110**
 119. **Figure 111**
 120. **Figure 112**
 121. **Figure 113**
 122. **Figure 114**
 123. **Figure 115**
 124. **Figure 116**
 125. **Figure 117**
 126. **Figure 118**
 127. **Figure 119**
 128. **Figure 120**
 129. **Figure 121**
 130. **Figure 122**
 131. **Figure 123**
 132. **Figure 124**
 133. **Figure 125**
 134. **Figure 126**
 135. **Figure 127**
 136. **Figure 128**
 137. **Figure 129**
 138. **Figure 130**
 139. **Figure 131**
 140. **Figure 132**
 141. **Figure 133**
 142. **Figure 134**
 143. **Figure 135**
 144. **Figure 136**
 145. **Figure 137**
 146. **Figure 138**
 147. **Figure 139**
 148. **Figure 140**
 149. **Figure 141**
 150. **Figure 142**
 151. **Figure 143**
 152. **Figure 144**
 153. **Figure 145**
 154. **Figure 146**
 155. **Figure 147**
 156. **Figure 148**
 157. **Figure 149**
 158. **Figure 150**
 159. **Figure 151**
 160. **Figure 152**
 161. **Figure 153**
 162. **Figure 154**
 163. **Figure 155**
 164. **Figure 156**
 165. **Figure 157**
 166. **Figure 158**
 167. **Figure 159**
 168. **Figure 160**
 169. **Figure 161**
 170. **Figure 162**
 171. **Figure 163**
 172. **Figure 164**
 173. **Figure 165**
 174. **Figure 166**
 175. **Figure 167**
 176. **Figure 168**
 177. **Figure 169**
 178. **Figure 170**
 179. **Figure 171**
 180. **Figure 172**
 181. **Figure 173**
 182. **Figure 174**
 183. **Figure 175**
 184. **Figure 176**
 185. **Figure 177**
 186. **Figure 178**
 187. **Figure 179**
 188. **Figure 180**
 189. **Figure 181**
 190. **Figure 182**
 191. **Figure 183**
 192. **Figure 184**
 193. **Figure 185**
 194. **Figure 186**
 195. **Figure 187**
 196. **Figure 188**
 197. **Figure 189**
 198. **Figure 190**
 199. **Figure 191**
 200. **Figure 192**
 201. **Figure 193**
 202. **Figure 194**
 203. **Figure 195**
 204. **Figure 196**
 205. **Figure 197**
 206. **Figure 198**
 207. **Figure 199**
 208. **Figure 200**
 209. **Figure 201**
 210. **Figure 202**
 211. **Figure 203**
 212. **Figure 204**
 213. **Figure 205**
 214. **Figure 206**
 215. **Figure 207**
 216. **Figure 208**
 217. **Figure 209**

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to hypothetical protein FLJ31614.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RC Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC023566; AAH23566.1; -
 DR F1R; PT0272; PT0271.
 DR GO: GO:000509; F:Calcium ion binding; IEA.
 DR GO: GO:000525; F:GTP binding; IEA.
 DR GO: GO:0003928; F:RAB small monomeric GTPase activity; IEA.
 DR GO: GO:0003930; F:RAS small monomeric GTPase activity; IEA.
 DR GO: GO:0003931; F:Rho small monomeric GTPase activity; IEA.
 DR GO: GO:0006886; P:intracellular protein transport; IEA.
 DR GO: GO:0007264; P:small GTPase mediated signal transduction; IEA.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR003577; GTPase_Ras.
 DR InterPro: IPR003578; GTPase_Rho.
 DR InterPro: IPR002041; RAN.
 DR InterPro: IPR001806; Ras transfrmng.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00036; efhand; 2.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR SMART: SM00054; EPH; 2.
 DR SMART: SM00175; RAB; 1.
 DR SMART: SM00176; RAN; 1.
 DR SMART: SM00173; RAS; 1.
 DR SMART: SM00174; RHO; 1.
 DR TIGRFAMs: TIGR00231; small_GTP; 1.
 DR PROSITE: PS00018; EF_HAND; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 740 AA; 82879 MW; 89AFCF4C159760F0 CRC64;

 Query Match 39.2%; Score 49; DB 4; Length 740;
 Best Local Similarity 52.2%; Pred. No. 59;
 Matches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

 QY 2 LGPFRSARRRRRFGDPTPKVLE 24
 DB 72 LGSIRGRRRDGFLDPAPAVSE 94

 RESULT 13
 ID Q8SYA1 PRELIMINARY; PRT; 947 AA.
 AC Q8SYA1
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE RH10407p.
 GN CG7097.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkley;
 RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Faise B.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY071688; AA149310.1; -
 DR FlyBase; FBgn0034421; CG7097.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004574; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR001180; Citron.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_Chir_kinase.
 DR Pfam: PF00780; CNH; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00036; CNH; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 947 AA; 105224 MW; 814262CDB95FF56D CRC64;

 Query Match 39.2%; Score 49; DB 5; Length 947;
 Best Local Similarity 52.2%; Pred. No. 76;
 Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

 QY 6 RSARRRREP-----GLDPTPKV 22
 DB 552 RSKRRHTPRPISNGLPPTPKV 574

 RESULT 14
 ID Q8MLI8 PRELIMINARY; PRT; 947 AA.
 AC Q8MLI8
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE CG7097-FB.
 GN CG7097.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne W., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paclos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbalw C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Wang X.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RA "Sequencing of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Masra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003796; AM70845.1; ..
DR FlyBase; FBgn0034421; CG7097.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR02230; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00219; TYK; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 947 AA; 105195 MW; 480EC84BAE0D020 CRC64;

Query Match

39.2%; Score 49; DB 5; Length 947;

Best Local Similarity 52.2%; Pred. No. 76;
Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;
QY 6 RSARRRRREP-----GLDPTPKV 22
DB 552 RSHKRHRTPRPISGNGLPPTPKV 574
RESULT 15
QV98R6
ID QV98R6 PRELIMINARY; PRT; 1218 AA.
AC QV98R6;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE CG7097 protein.
GN CG7097.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolchakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Foele C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN EMBL; AE003796; AM75595.1; ..
DR HSP; Q63450; IA06.
DR FlyBase; FBgn0034421; CG7097.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.

Tue Sep 21 14:48:45 2004

```
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00780; CNH; 1.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00036; CNH; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 1218 AA; 132395 MW; 78A7FAF880CEDC8 CRC64;

Query Match      39.2%; Score 49; DB 5; Length 1218;
Best Local Similarity 52.2%; Pred. No. 98;
Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

QY      6 RSARRRRP-----GLDPTPKV 22
      |||:|||||
DB     823 RSHKRRHTPPRFISNGLPPTPKV 845

Search completed: September 21, 2004, 13:03:45
Job time : 6.62011 secs
```

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 21, 2004, 12:43:42 ; Search time 4.57542 Seconds
(without alignments)
1482.081 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFRRSARRRRFGLDTPKYLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	24	AAB30927	Aab30927 Peptide d
2	125	100.0	1120	ABR58365	AbR58365 Human NOV
3	125	100.0	1188	AB30929	Ab30929 Amino aci
4	125	100.0	1188	AAU14467	Aau14467 Human nov
5	125	100.0	1188	AAU14231	Aau14231 Human nov
6	125	100.0	1188	AAU10551	Aau10551 Human A25
7	125	100.0	1188	AAU10551	Aau10551 Human A25
8	125	100.0	1188	ABR58365	AbR58365 Human NOV
9	125	100.0	1189	AB30929	Ab30929 Amino aci
10	125	100.0	1189	AAU14467	Aau14467 Human nov
11	125	100.0	1189	AAU14231	Aau14231 Human nov
12	125	100.0	1189	AAU10551	Aau10551 Human A25
13	125	100.0	1189	AAU10551	Aau10551 Human A25
14	125	100.0	1189	ABR58365	AbR58365 Human NOV
15	93	74.4	1188	AAU10551	Aau10551 Human A25
16	93	74.4	1188	AAU10551	Aau10551 Human A25
17	92	73.6	545	ABR72288	AbB72288 Murine pr
18	92	73.6	688	ABR72300	AbB72300 Rat prote
19	75	60.0	15	AA30928	Aa30928 Antigenic
20	52	41.6	956	AA30928	Aa30928 Rat Prote
21	51	40.8	161	AA30928	Aa30928 Antigenic
22	51	40.8	161	AA30928	Aa30928 Rat Prote
23	50	40.0	1024	AB30929	Ab30929 Amino aci
24	49	39.2	189	AB30929	Ab30929 Amino aci
25	49	39.2	466	AAU58625	Aau58625 Propionib

26	49	39.2	466	6	ABM55144	Abm55144 Propionib
27	49	39.2	616	6	ABM55406	Abm55406 Propionib
28	49	39.2	832	3	AAU58663	Aau58663 Human tyr
29	49	39.2	832	5	AAU71093	Aau71093 Human can
30	49	39.2	1218	4	ABR582736	AbR582736 Drosophil
31	48	38.4	100	4	AAO03386	Aao03386 Human pol
32	48	38.4	373	2	AAW24249	Aaw24249 Aquifex a
33	48	38.4	373	2	AAW24249	Aaw24249 Aquifex a
34	48	38.4	3588	2	AAU34712	Aau34712 Bacillus
35	47	37.6	69	4	AAU58901	Aau58901 Propionib
36	47	37.6	69	4	AAU58901	Aau58901 Propionib
37	47	37.6	69	6	ABM5420	Abm5420 Propionib
38	47	37.6	69	6	ABM5420	Abm5420 Propionib
39	47	37.6	72	2	AAW27427	Aaw27427 Human CRA
40	47	37.6	614	4	ABG24990	Abg24990 Novel hum
41	47	37.6	984	6	ABG24990	Abg24990 Novel hum
42	47	37.6	1245	4	ABG24990	Abg24990 Novel hum
43	46	36.8	99	4	AAU40886	Aau40886 Propionib
44	46	36.8	99	6	ABM37405	Abm37405 Propionib
45	46	36.8	114	4	AAU57162	Aau57162 Propionib

ALIGNMENTS

RESULT 1
AAB30927
ID AAB30927 standard; peptide; 24 AA.
XX
AC AAB30927;
XX
DT 02-APR-2001 (first entry)
XX
DE Peptide derived from a human alpha11 integrin chain.

XX Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
XX osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
XX wound healing; trauma; rheumatoid arthritis; osteoarthritis;
XX osteoporosis; cartilage damage; bone damage; cartilage.
XX Homo sapiens.
XX WO200075187-A1.
XX PD 14-DEC-2000.
XX PF 31-MAY-2000; 2000WO-SE001135.
XX PR 03-JUN-1999; 99SE-00002056.
XX (ACTI-) ACTIVE BIOTECH AB.
PI Gullberg D;
XX WPI; 2001-071061/08.
DR Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.
XX Claim 23; Page 40; 79pp; English.

XX The present sequence is derived from the cytoplasmic domain of the human
XX integrin subunit, designated alpha11. The alpha11 polynucleotide and
XX polypeptide are useful as markers of cell target molecules, such as
XX fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
XX derived cells or stem cells. They are also used for determining the
XX differential stage of cells during differentiation, development in
XX pathological conditions, in tissue regeneration, in transplantation or in
XX therapeutic and physiological repair of tissues. The pathological
XX conditions involving subunit alpha11 are selected from damage of cells,
XX muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
XX osteoarthritis and osteoporosis, damage of cartilage and bone, and

CC cartilage and bone diseases. The polypeptide is useful for detecting the
 CC formation of cartilage during embryonic development, for detecting
 CC physiological therapeutic repair of cartilage and muscle, for selection
 CC and analysis, or for sorting, isolating or purification of chondrocytes
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
 CC during transplantation of cartilage or chondrocytes during
 CC transplantation of cartilage or chondrocytes, respectively, or of muscle
 CC or muscle cells during transplantation of muscle or muscle cells,
 CC respectively, and for studies of differentiation or chondrocytes or
 CC muscle cells

XX Sequence 24 AA;

Query Match 100.0%; Score 125; DB 4; Length 24;
 Best Local Similarity 100.0%; Pred. NO. 1.3e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRPGLDPTPKVLE 24
 |||||
 Db 1 KLGFRRSARRRRPGLDPTPKVLE 24
 |||||

RESULT 2

ID ABR58365 standard; protein; 1120 AA.
 AC ABR58365;
 XX AC

07-JUN-2003 (first entry)

Human NOV2b.

Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;
 diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 neurodegenerative disorder; Alzheimer's disease; immune disorder;
 haematopoietic disorder.

XX Homo sapiens.

XX OS

XX PN W02003029423-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031358.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327342P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 29-OCT-2001; 2001US-0343629P.

XX PR 01-NOV-2001; 2001US-0349575P.

XX PR 12-APR-2002; 2002US-0346357P.

XX PR 12-APR-2002; 2002US-0371972P.

XX PR 17-APR-2002; 2002US-0371980P.

XX PR 19-APR-2002; 2002US-0373805P.

XX PR 23-APR-2002; 2002US-0374738P.

XX PR 16-MAY-2002; 2002US-0381101P.

XX PR 17-MAY-2002; 2002US-0381635P.

XX PR 29-MAY-2002; 2002US-0383830P.

XX PR 01-OCT-2002; 2002US-00262839.

XX PA (CURA-) CURAGEN CORP.

XX XX

PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;
 PI Kekuda R, Leach MD, Li L, Miller CE, Pattarajan M, Rieger DK;
 PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
 XX WPI; 2003-381625/36.
 DR N-PSDB; ACC72077.
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
 PT dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 PS Claim 1; Page 107; 487pp; English.
 XX The present invention relates to novel human NOV proteins and their
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
 CC proteins are useful in manufacturing a medicament for treating a syndrome
 CC associated with a human disease. The NOV proteins and coding sequences
 CC may be used to diagnose, treat or prevent metabolic disorders such as
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune
 CC disorders, haematopoietic disorders and various dyslipidaemias
 XX SQ Sequence 1120 AA;

Query Match 100.0%; Score 125; DB 6; Length 1120;
 Best Local Similarity 100.0%; Pred. NO. 6.2e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRPGLDPTPKVLE 24
 |||||
 Db 1097 KLGFRRSARRRRPGLDPTPKVLE 1120

RESULT 3

AAB30929

ID AAB30929 standard; protein; 1188 AA.

XX AAB30929;

XX 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha11 integrin chain.

XX Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
 KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
 KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
 KW osteoporosis; cartilage damage; bone damage; cartilage.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22 /note= "signal peptide"

FT Region 951..972 /note= "leucine zipper"

FT Domain 1142..1164 /note= "transmembrane domain"

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SB001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX XX

DR WPI; 2001-071061/08.
DR N-PSDB; AAC6871.
XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.
XX
PS Disclosure; Fig 2a-c; 79pp; English.
XX
CC The present sequence represents a human integrin subunit, designated
CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers
CC of cell target molecules, such as fibroblasts, muscle cells,
CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.
CC They are also used for determining the differential-stage of cells during
CC differentiation, development in pathological conditions, in tissue
CC regeneration, in transplantation or in therapeutic and physiological
CC repair of tissues. The pathological conditions involving subunit alpha11
CC are selected from damage of cells, muscle dystrophy, fibrosis, wound
CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,
CC damage of cartilage and bone, and cartilage and bone diseases. The
CC polypeptide is useful for detecting the formation of cartilage during
CC embryonic development, for detecting physiological therapeutic repair of
CC cartilage and muscle, for selection and analysis, or for sorting,
CC isolating or purification of chondrocytes and muscle cells, for detecting
CC regeneration of cartilage or chondrocytes during transplantation of
CC cartilage or chondrocytes during transplantation of cartilage or
CC chondrocytes, respectively, or of muscle or muscle cells during
CC transplantation of muscle or muscle cells, respectively, and for studies
CC of differentiation or chondrocytes or muscle cells
XX
SQ Sequence 1188 AA;
Query Match 100.0%; Score 125; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLGFPSARRRRPGLDTPKVL 24
DB 1165 KLGFPSARRRRPGLDTPKVL 1188
RESULT 4
AAU14467
XX AAU14467 standard; protein; 1188 AA.
AC AAU14467;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #338.
XX
KW Human; novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US002623.
XX
XX 25-JAN-2000; 2000US-00491404.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX

DR WPI; 2001-451939/48.
DR N-PSDB; AAS22772.
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage.
XX
PS Example 4; Page 828-831; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human proteins or
CC their active domains. The polypeptides, polynucleotides and antibodies
CC raised against the polypeptides are used in a method of treatment of a
CC mammal and prevention of disorders caused by the aberrant protein
CC expression or activity. The polypeptides can be used as molecular weight
CC markers, food supplements, and in antibody production. The polypeptides
CC are used to identify compounds which bind to the polypeptides.
CC Polynucleotides of the invention are used as probes and primers, for
CC sequencing, for chromosome or gene mapping, in the production of
CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
CC therapy. Polypeptides of the invention can be used to target drugs to a
CC tumour, in assays to determine biological activity, to raise
CC antibodies/elicit an immune response, to determine quantitative protein
CC levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
CC diseases, nervous system disorders, and infection. The present sequence
CC represents a protein of the invention
XX
SQ Sequence 1188 AA;
Query Match 100.0%; Score 125; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLGFPSARRRRPGLDTPKVL 24
DB 1165 KLGFPSARRRRPGLDTPKVL 1188
RESULT 5
AAU14231
XX AAU14231 standard; protein; 1188 AA.
AC AAU14231;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #102.
XX
KW Human; novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US002623.
XX
XX 25-JAN-2000; 2000US-00491404.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX

PA (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-451939/48.
 DR N-PSDB; AAS22536.
 XX
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage.
 PT
 XX
 PS Example 4; Page 578-581; 894pp; English.
 XX
 CC The invention relates to polynucleotides encoding novel human proteins or
 CC their active domains. The polypeptides, polynucleotides and antibodies
 CC raised against the polypeptides are used in a method of treatment of a
 CC mammal and prevention of disorders caused by the aberrant protein
 CC expression or activity. The polypeptides can be used as molecular weight
 CC markers, food supplements, and in antibody production. The polypeptides
 CC are used to identify compounds which bind to the polypeptides.
 CC Polynucleotides of the invention are used as probes and primers, for
 CC sequencing, for chromosome or gene mapping, in the production of
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
 CC therapy. Polypeptides of the invention can be used to target drugs to a
 CC tumour, in assays to determine biological activity, to raise
 CC antibodies/ elicit an immune response, to determine quantitative protein
 CC levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
 CC diseases, nervous system disorders, and infection. The present sequence
 CC represents a protein of the invention
 XX
 SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KLGFFRSARRRPGDPTPKVLE 24
 Db 1165 KLGFFRSARRRPGDPTPKVLE 1188
 RESULT 6
 AAB50085
 ID AAB50085 standard; protein; 1188 AA.
 AC AAB50085;
 XX
 XX 19-MAR-2001 (first entry)
 DT
 DE Human A259.
 XX
 KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
 KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
 KW rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..1141 /label= Extracellular_domain
 FT Peptide 1..22 /label= Signal_peptide
 FT Protein 22..1188 /label= Mature_protein
 FT Domain 39..174

FT
 FT /label= Integrin_alphasubunit_repeat_domain_#1
 FT 115..157
 FT /label= Integrin_alphasubunit_repeat_domain_#2
 FT 164..345
 FT /label= I_domain
 FT 367..392
 FT /label= Integrin_alphasubunit_repeat_domain_#3
 FT 421..455
 FT /label= Integrin_alphasubunit_repeat_domain_#4
 FT 478..516
 FT /label= Integrin_alphasubunit_repeat_domain_#5
 FT 540..575
 FT /label= Integrin_alphasubunit_repeat_domain_#6
 FT 602..640
 FT /label= Integrin_alphasubunit_repeat_domain_#7
 FT 1142..1164
 FT /label= Transmembrane_domain
 FT 1165..1188
 FT /label= Cytoplasmic_domain
 XX
 PN WO200073339-A1.
 XX
 XX 07-DEC-2000.
 XX
 XX 15-MAY-2000; 2000WO-US013262.
 XX
 XX 28-MAY-1999; 99US-00322790.
 XX
 XX 27-APR-2000; 2000US-00561263.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Pan Y, Lora JM;
 XX
 XX WPI; 2001-041142/05.
 DR
 DR N-PSDB; AAC91901, AAC91902.
 XX
 PT Nucleic acid encoding alpha-integrin subunits, useful for treatment and
 PT diagnosis of fibrosis, e.g. of the liver.
 XX
 PS Claim 8; Fig 1; 164pp; English.
 XX
 CC The present sequence is human integrin alpha subunit, A259. A259 is
 CC homologous with the alpha1 and alpha10 integrin subunits and is
 CC overexpressed in fibrosis. A259 is implicated in regulation of
 CC proliferation, differentiation and/or function of many different cell
 CC types. Inhibitors of A259 activity are useful for the treatment of liver
 CC disease, particularly fibrosis, and also fibrosis in other organs.
 CC (specifically lung and kidney). In addition, A259 can be used for
 CC treatment and prevention of cancer, osteoporosis, acute myeloid
 CC leukaemia, HIV infection, and rheumatoid arthritis
 XX
 SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KLGFFRSARRRPGDPTPKVLE 24
 Db 1165 KLGFFRSARRRPGDPTPKVLE 1188
 RESULT 7
 AAU10551
 ID AAU10551 standard; protein; 1188 AA.
 XX
 XX AAU10551;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 XX Human A259 polypeptide.
 XX
 KW Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein;

XX SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 7; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRSARRRRPGLDPTPKVLE 24
 |||||
 DB 1165 KLGFRSARRRRPGLDPTPKVLE 1188

RESULT 9
 AAB25582
 ID AAB25582 standard; protein; 1189 AA.
 XX AC AAB25582;
 XX DT 21-NOV-2000 (first entry)
 XX DE ITGA11 protein encoded by human secreted protein gene #7.
 XX KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
 XX OS Homo sapiens.
 XX PN WO200029435-A1.
 XX PD 25-MAY-2000.
 XX PF 27-OCT-1999; 99WO-US025031.
 XX PR 28-OCT-1998; 98US-0105971P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JU, Moore PA, Wei Y;
 PI Greene JW;
 XX WPI; 2000-387742/33.
 XX N-PSDB; AAA80612.
 XX PT Isolated nucleic acid molecules encoding human secreted proteins are used
 PT for the prevention, amelioration and treatment of autoimmune,
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 PT wounds, and infectious diseases.
 XX PS Claim 1; Fig 19A-F; 803pp; English.

CC The present invention relates to 12 secreted human proteins and the
 CC nucleotide sequences encoding them. The polynucleotide sequences given in
 CC AAA80606-A80623 encode the 12 secreted protein sequences given in
 CC AAB25576-B25593. The human secreted proteins have various activities
 CC dependent on the tissues in which they are expressed. Examples of the
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;
 CC antiarthritic; antirheumatic; dermatological; antiproliferative;
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;
 CC and antifungal activity. The proteins, polypeptides, agonists and
 CC antagonists may be used to treat prevent and/or diagnose various disease,
 CC disorders and conditions examples of which include: immune disorders e.g.
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
 CC Crohn's disease and nephritis; hyperproliferative disorders such as
 CC paraneoplasias and purpura; cardiovascular disorders e.g. coronary
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
 CC proteins and polynucleotide sequences may also be used in wound healing

CC and the treatment of infectious diseases. The human secreted protein gene
 CC #7 and protein sequences are represented in sequences AAA80612 and
 CC AAB25582. Secreted protein gene #7 is located at position chromosome 15
 CC q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
 CC the secreted protein gene#7

XX SQ Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 3; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRSARRRRPGLDPTPKVLE 24
 |||||
 DB 1166 KLGFRSARRRRPGLDPTPKVLE 1189

RESULT 10
 ABG12949
 ID ABG12949 standard; protein; 1189 AA.
 XX AC ABG12949;
 XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #12940.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS77136.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 43308; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 4; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREGLDTPPKVLE 24
DB 1166 KLGFFRSARRRREGLDTPPKVLE 1189

RESULT 11
ABR58364
ID ABR58364 standard; protein; 1189 AA.
XX
AC ABR58364;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human NOV2a.
XX
KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
KW antiparkinsonian; antihypertensive; gene therapy; metabolic disorder;
KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
KW haematopoietic disorder.
XX
OS Homo sapiens.
XX
PN WO2003029423-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031358.
XX
PR 02-OCT-2001; 2001US-0326483P.
XX
PR 03-OCT-2001; 2001US-0327342P.
XX
PR 03-OCT-2001; 2001US-0327917P.
XX
PR 03-OCT-2001; 2001US-0328029P.
XX
PR 03-OCT-2001; 2001US-0328044P.
XX
PR 03-OCT-2001; 2001US-0328056P.
XX
PR 12-OCT-2001; 2001US-0328849P.
XX
PR 15-OCT-2001; 2001US-0329414P.
XX
PR 17-OCT-2001; 2001US-0330142P.
XX
PR 22-OCT-2001; 2001US-0341658P.
XX
PR 24-OCT-2001; 2001US-0339286P.
XX
PR 24-OCT-2001; 2001US-0343629P.
XX
PR 29-OCT-2001; 2001US-0349575P.
XX
PR 01-NOV-2001; 2001US-0346357P.
XX
PR 12-APR-2002; 2002US-0371972P.
XX
PR 12-APR-2002; 2002US-0371980P.
XX
PR 17-APR-2002; 2002US-0373261P.
XX
PR 19-APR-2002; 2002US-0373805P.
XX
PR 23-APR-2002; 2002US-0374738P.
XX
PR 16-MAY-2002; 2002US-0381101P.
XX
PR 17-MAY-2002; 2002US-0381635P.
XX
PR 29-MAY-2002; 2002US-0383830P.
XX
PR 01-OCT-2002; 2002US-00262839.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
PI Edinger SA, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;
PI Rothenberg ME, Shinkens RA, Smithson G, Spytek KA, Taupier RJ;
PI Vernet CAM, Voss EZ, Zerhusen ED, Zhong M;
XX
WPI; 2003-381625/36.

DR N-PSDB; ACC72076.
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
PT dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 105; 487pp; English.
XX
CC The present invention relates to novel human NOV proteins and their
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
CC proteins are useful in manufacturing a medicament for treating a syndrome
CC associated with a human disease. The NOV proteins and coding sequences
CC may be used to diagnose, treat or prevent metabolic disorders such as
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
CC disorders such as Alzheimer's disease or Parkinson's disease, immune
CC disorders, haematopoietic disorders and various dyslipidaemias
XX
SQ Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 6; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREGLDTPPKVLE 24
DB 1166 KLGFFRSARRRREGLDTPPKVLE 1189

RESULT 12
ADA27054
ID ADA27054 standard; protein; 1189 AA.
XX
AC ADA27054;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human novel secreted protein from cDNA HOHEY69 #1.
XX
KW cytostatic; antiinflammatory; immunomodulator; neuroprotective;
KW hemostatic; gene therapy; cancer; inflammation; immune disorder;
KW neurological disorder; blood clotting disorder; food additive;
KW preservative; human; secreted protein.
XX
OS Homo sapiens.
XX
PN US2003055231-A1.
XX
PD 20-MAR-2003.
XX
PF 29-OCT-2001; 2001US-00984130.
XX
PR 28-OCT-1998; 98US-0105971P.
XX
PR 27-OCT-1999; 99WO-US025031.
XX
PR 19-APR-2000; 2000US-0198407P.
XX
PR 30-OCT-2000; 2000US-0243792P.
XX
PR 18-APR-2001; 2001US-00836353.
XX
PA (NIJ/) NI J.
PA (YOUN/) YOUNG P E.
PA (KERN/) KENNY J J.
PA (OLSE/) OLSEN H S.
PA (MOOR/) MOORE P A.
PA (WEIY/) WEI Y.
PA (GREE/) GREENE J M.
PA (RUBE/) RUBEN S M.
PA (LIUD/) LIU D.
PA (CROC/) CROCKER P R.
XX
PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM, Liu D, Crocker PR;
XX
WPI; 2003-567103/53.

DR N-PSDB; ADA27036.

XX New human secreted nucleic acid molecules and polypeptides, useful for

XX preventing, treating, or ameliorating a medical condition, such as

PT cancer, inflammation, immune disorders, neurological and blood clotting

PT disorders.

XX

PS Claim 11; Fig 19; 454pp; English.

XX

CC The invention relates to an isolated nucleic molecule that is at least

CC 95% identical to 18 human cDNA sequences representing 12 novel genes

CC encoding secreted proteins or a polynucleotide fragment of the cDNA

CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic

CC variant of the polynucleotide having a polynucleotide capable of

CC hybridising under conditions the polynucleotide, where the polynucleotide

CC having a nucleotide sequence of only A or T residues. Also included are

CC recombinant vectors, host cells (for producing the polypeptide), the

CC secreted polypeptide (comprising a sequence that is at least 95%

CC identical to a polypeptide fragment, domain, epitope, full-length

CC protein, variant, allelic variant or species homologue), antibodies that

CC specifically bind to the polypeptides, diagnosing, treating, preventing

CC or ameliorating a medical condition by administering the polynucleotide

CC or the polypeptide, the gene corresponding to the cDNA sequence and

CC identifying an activity in a biological assay (by expressing the cDNA

CC sequence in a cell, isolating the supernatant, and detecting an activity

CC in a biological assay and identifying the protein in the supernatant

CC having the activity). The polypeptides, nucleic acids and antibodies are

CC useful for diagnosing a pathological condition or a susceptibility to a

CC pathological condition, for preventing, treating, or ameliorating a

CC medical condition, such as cancer, inflammation and other immune

CC disorders, neurological and blood clotting disorders (many examples

CC given in the specification). The nucleic acids are also useful for

CC chromosome identification, radiation hybrid mapping or long-range

CC restriction mapping. The polypeptides and antibodies are useful for

CC providing immunological probes for differential identification of the

CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,

CC agonist or antagonist may also be used as a food additive or preservative

CC to increase or decrease storage capabilities, fat content or other

CC nutritional components. The present is a secreted protein of the

CC invention.

XX

Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 6; Length 1189;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRPGDPTPKVLE 24

Db 1166 KLGFFSARRRRPGDPTPKVLE 1189

RESULT 13

AD63570

ID ADE63570 standard; protein; 1189 AA.

XX ADE63570;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein Q9URX5, SEQ ID NO 9514.

XX

XX Human; pain; neuronal tissue; gene therapy;

XX Spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

FN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

BR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-269312/26.

XX

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page: 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 7; Length 1189;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRPGDPTPKVLE 24

Db 1166 KLGFFSARRRRPGDPTPKVLE 1189

RESULT 14

AD63584

ID ADE86584 standard; protein; 1189 AA.

XX ADE86584;

XX

DT 29-JAN-2004 (first entry)

XX

DE Novel human secreted protein #7.

XX

XX human; secreted protein; cancer; liver disorder; hepatitis;

XX neural disorder; Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN US2003129685-A1.
XX 10-JUL-2003.
XX 18-APR-2001; 2001US-00836353.
XX 28-OCT-1998; 98US-0105971P.
PR 27-OCT-1999; 99WO-US025031.
PR 19-APR-2000; 2000US-0198407P.
XX (NIJ/) NI J.
PA (YOUNG/) YOUNG P E.
PA (KERN/) KENNY J J.
PA (OLSEN/) OLSEN H S.
PA (MOORE/) MOORE P A.
PA (WEI/) WEI Y.
PA (GREENE/) GREENE J M.
PA (RUBEN/) RUBEN S M.
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM;
XX WPI; 2004-020335/02.
DR N-PSDB; ADB86566.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX Claim 11; SEQ ID NO 35; 380pp; English.
XX The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of a novel human secreted protein.
XX Sequence 1189 AA;
Query Match 100.0%; Score 125; DB 8; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLGFFRSARRRREPGLDTPKVL 24
Db 1166 KLGFFRSARRRREPGLDTPKVL 1189
RESULT 15
AAB50087
ID AAB50087 standard; protein; 1188 AA.
XX AAB50087;
AC AC
XX 19-MAR-2001 (first entry)
DT DT
XX Murine A259.
DE DE
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX Mus sp.
OS OS
XX Location/Qualifiers
FH Key
FT Domain 1..1141
FT /label= Extracellular_domain
FT Peptide 1..22
FT /label= Signal_peptide
FT Protein 23..1188
FT /label= Mature_protein

FT Domain 39..74
FT /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain 115..157
FT /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain 164..345
FT /label= I domain
FT Domain 367..392
FT /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain 421..455
FT /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain 478..516
FT /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain 540..575
FT /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain 602..640
FT /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain 1142..1164
FT /label= Transmembrane_domain
FT Domain 1165..1188
FT /label= Cytoplasmic_domain
XX WO200073339-A1.
XX 07-DEC-2000.
XX 15-MAY-2000; 2000WO-US013262.
XX 28-MAY-1999; 99US-00322790.
PR 27-APR-2000; 2000US-00561263.
XX (MILL-) MILLENNIUM PHARM INC.
XX Pan Y, Lora JM;
XX WPI; 2001-041142/05
XX N-PSDB; AAC91904; AAC91905.
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
PT diagnosis of fibrosis, e.g. of the liver.
XX Claim 8; Fig 5; 164pp; English.
XX The present sequence is murine integrin alpha subunit, A259. A259 is
CC homologous with the alpha and alpha0 integrin subunits and is
CC overexpressed in fibrosis. A259 is implicated in regulation of
CC proliferation, differentiation and/or function of many different cell
CC types. Inhibitors of A259 activity are useful for the treatment of liver
CC disease, particularly fibrosis, and also fibrosis in other organs
CC (specifically lung and kidney). In addition, A259 can be used for
CC treatment and prevention of cancer, osteoporosis, acute myeloid
CC leukaemia, HIV infection, and rheumatoid arthritis
XX Sequence 1188 AA;
Query Match 74.4%; Score 93; DB 4; Length 1188;
Best Local Similarity 70.8%; Pred. No. 0.00022;
Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KLGFFRSARRRREPGLDTPKVL 24
Db 1165 KLGFFRSARRRREPGLDTPKVL 1188

Search completed: September 21, 2004, 12:59:13
Job time : 6.57542 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:03:59 ; Search time 4.32402 Seconds
(without alignments)
1782.414 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFRRSARRRRPGLDTPKYLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	125	100.0	1120	12	US-10-262-839-6 Sequence 6, Appli
2	125	100.0	1188	15	US-10-231-263-338 Sequence 338, App
3	125	100.0	1188	15	US-10-291-263-810 Sequence 810, App
4	125	100.0	1189	10	US-09-984-130-35 Sequence 35, Appl
5	125	100.0	1189	10	US-09-836-353A-35 Sequence 35, Appl
6	125	100.0	1189	12	US-10-262-839-4 Sequence 4, Appli
7	92	73.6	545	10	US-09-866-050A-500 Sequence 500, App
8	92	73.6	688	10	US-09-866-050A-624 Sequence 624, App
9	52	41.6	420	16	US-10-437-963-106810 Sequence 106810,
10	51.5	41.2	2614	16	US-10-437-963-188563 Sequence 188563,
11	51	40.8	161	10	US-09-764-891-5091 Sequence 5091, Ap
12	51	40.8	161	14	US-10-205-428-457 Sequence 457, App
13	50	40.0	68	16	US-10-437-963-125024 Sequence 125024,
14	50	40.0	115	12	US-10-425-114-48965 Sequence 48965, A
15	50	40.0	117	12	US-10-424-599-170608 Sequence 170608,

16	50	40.0	231	12	US-10-425-114-40558 Sequence 40558, A
17	50	40.0	767	16	US-10-437-963-180056 Sequence 180056,
18	49.5	39.6	72	12	US-10-424-599-220139 Sequence 220139,
19	49.5	39.6	1316	16	US-10-437-963-185606 Sequence 185606,
20	49	39.2	161	16	US-10-437-963-127748 Sequence 127748,
21	49	39.2	189	15	US-10-094-749-2410 Sequence 2410, Ap
22	49	39.2	832	9	US-09-834-765-2 Sequence 2, Appli
23	49	39.2	2309	16	US-10-437-963-188616 Sequence 188616,
24	48	38.4	112	12	US-10-424-599-244523 Sequence 244523,
25	48	38.4	139	16	US-10-437-963-113972 Sequence 113972,
26	48	38.4	160	16	US-10-437-963-173508 Sequence 173508,
27	48	38.4	301	12	US-10-425-114-67432 Sequence 67432, A
28	48	38.4	373	9	US-09-905-173-26 Sequence 26, Appl
29	48	38.4	373	14	US-10-060-432-26 Sequence 26, Appl
30	48	38.4	373	15	US-10-369-493-30 Sequence 30, Appl
31	48	38.4	528	15	US-10-369-493-22651 Sequence 22651, A
32	48	38.4	837	16	US-10-437-963-114561 Sequence 114561,
33	48	38.4	894	16	US-10-437-963-156468 Sequence 156468,
34	48	38.4	1050	16	US-10-437-963-130197 Sequence 130197,
35	48	38.4	1494	16	US-10-437-963-194014 Sequence 194014,
36	47.5	38.0	624	16	US-10-437-963-183985 Sequence 183985,
37	47	37.6	72	16	US-10-437-963-113590 Sequence 113590,
38	47	37.6	81	12	US-10-424-599-232098 Sequence 232098,
39	47	37.6	109	16	US-10-767-701-48187 Sequence 48187, A
40	47	37.6	132	12	US-10-424-599-216264 Sequence 216264,
41	47	37.6	204	16	US-10-437-963-161117 Sequence 161117,
42	47	37.6	250	16	US-10-437-963-127750 Sequence 127750,
43	47	37.6	387	12	US-10-425-114-60791 Sequence 60791, A
44	47	37.6	398	12	US-10-425-114-49090 Sequence 49090, A
45	47	37.6	486	16	US-10-437-963-125679 Sequence 125679,

ALIGNMENTS

RESULT 1
US-10-262-839-6
; Sequence 6, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:

; APPLICANT: Alsobrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917

```
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 6
; LENGTH: 1120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-6
```

```
Query Match      100.0%; Score 125; DB 12; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRPGLDPTPKVLE 24
   |||||
Db 1097 KLGFFRSARRRRPGLDPTPKVLE 1120
```

```
RESULT 2
US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-338
```

```
Query Match      100.0%; Score 125; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRPGLDPTPKVLE 24
   |||||
Db 1165 KLGFFRSARRRRPGLDPTPKVLE 1188
```

```
RESULT 3
US-10-291-265-810
```

```
; Sequence 810, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 810
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-810
```

```
Query Match      100.0%; Score 125; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRPGLDPTPKVLE 24
   |||||
Db 1165 KLGFFRSARRRRPGLDPTPKVLE 1188
```

```
RESULT 4
US-09-984-130-35
; Sequence 35, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PP489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-35
```

```
Query Match      100.0%; Score 125; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRPGLDPTPKVLE 24
   |||||
Db 1166 KLGFFRSARRRRPGLDPTPKVLE 1189
```

```
RESULT 5
US-09-836-353A-35
```

; Sequence 35, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489PL
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-35

Query Match 100.0%; Score 125; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 1166 KLGFFRSARRRREPGLDPTPKVLE 1189

RESULT 6

US-10-262-839-4

; Sequence 4, Application US/10262839

; Publication No. US20040038877A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, John.
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc.
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elna,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenna,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-4

Query Match 100.0%; Score 125; DB 12; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 1166 KLGFFRSARRRREPGLDPTPKVLE 1189

RESULT 7

US-09-866-050A-500

; Sequence 500, Application US/09866050A

; Publication No. US20030040471A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-500

Query Match 73.6%; Score 92; DB 10; Length 545;
Best Local Similarity 70.8%; Pred. No. 9.2e-05;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 522 KLGFFRSARRRREPGLDPTPKVLE 545

RESULT 8

US-09-866-050A-624

; Sequence 624, Application US/09866050A

; Publication No. US20030040471A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 624
LENGTH: 688
TYPE: PRT
ORGANISM: Rat
US-09-866-050A-624

Query Match 73.6%; Score 92; DB 10; Length 688;
Best Local Similarity 70.8%; Pred. No. 0.00012;
Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 665 KLGFFSARRRRRPGLDPTPKVLE 688

RESULT 9
US-10-437-963-106810
Sequence 106810, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 106810
LENGTH: 420
TYPE: PRT
FEATURE:
ORGANISM: Oryza sativa
OTHER INFORMATION: Clone ID: PAT_MRT4530_11220C.1.pap
US-10-437-963-106810

Query Match 41.6%; Score 52; DB 16; Length 420;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 LGGFFSARRRRRPGLDPTPK 21
|||:||||:||||:||||:||||:
Db 73 LGGFFSARRRRRPGLDPTPK 92

RESULT 10
US-10-437-963-188563
Sequence 188563, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 188563
LENGTH: 2614
TYPE: PRT
ORGANISM: Oryza sativa
OTHER INFORMATION: Clone ID: PAT_MRT4530_8515SC.1.pap
US-10-437-963-188563

Query Match 41.2%; Score 51; DB 16; Length 2614;
Best Local Similarity 46.9%; Pred. No. 2.9e+02;
Matches 15; Conservative 2; Mismatches 6; Indels 9; Gaps 1;

QY 2 LGGFFSARRRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 959 LGGFFSARRRRRPGLDPTPKVLE 990

RESULT 11
US-09-764-891-5091
Sequence 5091, Application US/09764891
Publication No. US20030077809A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5091
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-5091

Query Match 40.8%; Score 51; DB 10; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 RSARRRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 35 RSARRRRRPGLDPTPKVLE 53

RESULT 12
US-10-205-428-457
Sequence 457, Application US/10205428

Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAIL7C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 457
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-205-428-457
Query Match 40.8%; Score 51; DB 14; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDPTPKVLE 24
Db 35 RSRRRRQPDLTSPGPPLE 53
RESULT 13
US-10-437-963-125024
; Sequence 125024, Application US/10437963
; Publication No. US20040123343A1
Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 125024
LENGTH: 68
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(68)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_27708C.1.pep
US-10-437-963-125024
Query Match 40.0%; Score 50; DB 16; Length 68;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 7 SARRRREPGLDPTPK 21
Db 29 SARRRRPPRRPTPR 43
RESULT 14
US-10-425-114-48965
; Sequence 48965, Application US/10425114
; Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 48965
LENGTH: 115
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4570-009-H8_FLI.pep
US-10-425-114-48965
Query Match 40.0%; Score 50; DB 12; Length 115;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDP 18
Db 72 RAARRRDFGVQP 84
RESULT 15
US-10-424-599-170608
; Sequence 170608, Application US/10424599

Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAIL7C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 457
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-205-428-457
Query Match 40.8%; Score 51; DB 14; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDPTPKVLE 24
Db 35 RSRRRRQPDLTSPGPPLE 53
RESULT 13
US-10-437-963-125024
; Sequence 125024, Application US/10437963
; Publication No. US20040123343A1

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170608
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(117)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125074C.1.pcp
US-10-424-599-170608

```

```

Query Match      40.0%; Score 50; DB 12; Length 117;
Best Local Similarity 58.8%; Pred. No. 20;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      4 FERSARRRPGLDPTP 20
      |||||
Db      50 FFRKRESHEPELAPPP 66

```

```

Search completed: September 21, 2004, 13:29:02
Job time : 5.32402 secs

```


Sequence 31102, A
Sequence 22868, A
Sequence 29989, A
Sequence 29455, A
Sequence 38, Appl
Sequence 38, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 16809, A
Sequence 1771, A
Sequence 19469, A

28 46 36.8 298 4 US-09-252-991A-31102
29 46 36.8 318 4 US-09-252-991A-22868
30 46 36.8 329 4 US-09-252-991A-29989
31 46 36.8 349 4 US-09-252-991A-29455
32 46 36.8 438 2 US-08-577-492-38
33 46 36.8 438 2 US-09-079-630-38
34 46 36.8 481 1 US-08-286-856C-2
35 46 36.8 481 1 US-08-472-831-2
36 46 36.8 501 3 US-09-079-630-40
37 46 36.8 501 3 US-09-079-630-40
38 46 36.8 506 1 US-08-286-856C-3
39 46 36.8 506 1 US-08-472-831-3
40 46 36.8 606 2 US-08-577-492-32
41 46 36.8 606 2 US-09-079-630-32
42 46 36.8 691 4 US-09-252-991A-16809
43 45.5 36.4 294 4 US-09-252-991A-32345
44 45 36.0 133 4 US-09-252-991A-1771
45 45 36.0 148 4 US-09-252-991A-19469

ALIGNMENTS

RESULT 1
US-09-252-991A-19842
; Sequence 19842, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19842
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19842

Query Match 43.2%; Score 54; DB 4; Length 353;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 6 RSARRRRPGDPTP 20
|:|||||:
Db 119 RAARRRRDGPALPAP 133
|:|||||:
RESULT 2
US-09-252-991A-19055
; Sequence 19055, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19055
; LENGTH: 424
; TYPE: PRT

389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	54	43.2	353	Sequence 19842, A
2	52	41.6	424	Sequence 19055, A
3	52	41.6	690	Sequence 24903, A
4	51	40.8	272	Sequence 31371, A
5	50	40.0	425	Sequence 26326, A
6	49	39.2	236	Sequence 30071, A
7	48	38.4	162	Sequence 24838, A
8	48	38.4	373	Sequence 26, Appl
9	48	38.4	373	Sequence 26, Appl
10	48	38.4	373	Sequence 26, Appl
11	48	38.4	373	Sequence 26, Appl
12	48	38.4	573	Sequence 22802, A
13	48	38.4	885	Sequence 26129, A
14	47	37.6	88	Sequence 6510, A
15	47	37.6	110	Sequence 17182, A
16	47	37.6	166	Sequence 21902, A
17	47	37.6	256	Sequence 29580, A
18	47	37.6	310	Sequence 23221, A
19	47	37.6	559	Sequence 24692, A
20	47	37.6	600	Sequence 24130, A
21	46	36.8	212	Sequence 31401, A
22	46	36.8	222	Sequence 46, Appl
23	46	36.8	222	Sequence 43, Appl
24	46	36.8	222	Sequence 54, Appl
25	46	36.8	223	Sequence 46, Appl
26	46	36.8	253	Sequence 46, Appl
27	46	36.8	253	Sequence 46, Appl

ORGANISM: Pseudomonas aeruginosa

FEATURE:
NAME/KEY: UNSURE

LOCATION: (176),(216)

OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

Query Match 41.6%; Score 52; DB 4; Length 424;

Best Local Similarity 66.7%; Pred. No. 4.2;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 RRRREPGLDTPKVL 23

|||||

56 RRRRPGDPLPRL 70

RESULT 3

US-09-252-991A-24903

; Sequence 24903, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 24903

; LENGTH: 690

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (140),(161)

OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-24903

Query Match 41.6%; Score 52; DB 4; Length 690;

Best Local Similarity 52.8%; Pred. No. 7.1;

Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 GFFRSARRRRREPGLDTPK 21

|||||

23 GTRRRRRRPRCLEDPE 41

RESULT 4

US-09-252-991A-31371

; Sequence 31371, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31371

; LENGTH: 272

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31371

Query Match 40.8%; Score 51; DB 4; Length 272;

Best Local Similarity 64.3%; Pred. No. 3.7;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SARRRREPGLDPTP 20

|||||

176 TARRRRPGASPNP 189

RESULT 5

US-09-252-991A-26326

; Sequence 26326, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26326

; LENGTH: 425

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26326

Query Match 40.0%; Score 50; DB 4; Length 425;

Best Local Similarity 61.1%; Pred. No. 8.5;

Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 6 RSARRRRREPGLDTPPKVL 23

|||||

15 RSAGRRRSPGLRQOPDLL 32

RESULT 6

US-09-252-991A-30071

; Sequence 30071, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30071

; LENGTH: 296

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30071

Query Match 39.2%; Score 49; DB 4; Length 296;

Best Local Similarity 64.3%; Pred. No. 8.2;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 10 RRRREPGLDTPPKVL 23

|||||

1 RRRDFDRDPGLL 14

RESULT 7

US-09-252-991A-24838

; Sequence 24838, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24838
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24838

Query Match 38.4%; Score 48; DB 4; Length 162;
Best Local Similarity 50.0%; Pred. No. 6.1;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTP 20
Db 86 RLAAFRPHRRPQGRGGTF 105

RESULT 8
US-09-599-171A-26
; Sequence 26, Application US/08599171A
; Patent No. 5814473
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-09-599-171A-26

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLPSPKWME 43

RESULT 9
US-08-646-590B-26
; Sequence 26, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-646-590B-26

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLPSPKWME 43

RESULT 10
US-09-069-226-26
; Sequence 26, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES

```
;
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,226
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/599,171
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; FRAGMENT TYPE:
; US-09-069-226-26

Query Match 38.4%; Score 48; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLESPKVMK 43

RESULT 11
US-09-412-184-26
; Sequence 26, Application US/09412184
; Patent No. 6268188
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,184
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US/08/646,590
; FILING DATE: 08-May-1996

;
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5039
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-412-184-26

Query Match 38.4%; Score 48; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLESPKVMK 43

RESULT 12
US-09-252-991A-22802
; Sequence 22802, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22802
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22802

Query Match 38.4%; Score 48; DB 4; Length 573;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RSARRRREPGLDPTTP 20
Db 358 RSARCQRWPGVEPVP 372

RESULT 13
US-09-252-991A-26129
; Sequence 26129, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

Query Match	37.6%	Score 47;	DB 4;	Length 110;
Best Local Similarity	60.0%	Pred. No. 5.7;		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:24:50 ; Search time 27 Seconds
(without alignments)
4232.427 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRRRGLDPTPKVLE 1188

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	1.3	272	2 A5348	integrin alpha-1
2	15	1.3	1151	2 A5226	integrin alpha-1 c
3	15	1.3	1180	2 A35854	integrin alpha-1 c
4	11	0.9	1170	2 I45914	integrin alpha 2 s
5	11	0.9	1181	2 A33998	integrin alpha-2 c
6	9	0.8	74	2 I51524	integrin alpha-2 s
7	9	0.8	315	2 A43567	permease [imported
8	9	0.8	371	1 T13407	3-isopropylmalate
9	9	0.8	607	2 S0658	legumin - Gnetum g
10	9	0.8	1178	2 S44142	VIA-2 protein homo
11	8	0.7	20	2 A60822	cytochrome P450 PB
12	8	0.7	42	2 A34259	cytochrome P450mt4
13	8	0.7	76	2 I51527	integrin alpha 5 s
14	8	0.7	76	2 A43537	heat-stable antige
15	8	0.7	76	2 I53107	CD24 precursor - r
16	8	0.7	80	2 A48996	B cell surface ant
17	8	0.7	103	2 A70856	conserved hypotet
18	8	0.7	127	2 B75301	hypothetical prote
19	8	0.7	141	2 S15785	heat-stable antige
20	8	0.7	224	2 E71228	hypothetical prote
21	8	0.7	226	2 S76800	hypothetical prote
22	8	0.7	234	1 S15102	eosinophil major b
23	8	0.7	264	2 S22090	catechol O-methyl
24	8	0.7	308	2 A26337	conserved hypotet
25	8	0.7	356	2 F97419	BH1459 conserved h
26	8	0.7	360	2 A85016	hypothetical prote
27	8	0.7	367	2 S19172	cytochrome P450 2B
28	8	0.7	387	1 DYH0D4	dopamine receptor
29	8	0.7	402	1 S23860	chloramphenicol re

30	0.7	415	2 G83568	probable permease
31	0.7	442	2 AG3504	dihydrofolate synt
32	0.7	444	2 T01721	hypothetical prote
33	0.7	487	2 T47107	benzaldehyde dehyd
34	0.7	491	1 O4RTPB	cytochrome P450 2B
35	0.7	491	1 O4RTF2	cytochrome P450 2B
36	0.7	500	2 B1047	testosterone 16alp
37	0.7	547	2 AC1314	malolactic enzyme
38	0.7	547	2 AC1686	malolactic enzyme
39	0.7	605	1 W1WLEB	El protein - bovin
40	0.7	606	2 A72429	oligopeptide ABC t
41	0.7	614	2 A69845	Na+/H+ antiporter
42	0.7	620	1 W1WLS2	El protein - bovin
43	0.7	685	2 AC0527	ferriochrome transp
44	0.7	697	2 T16306	hypothetical prote
45	0.7	739	2 A88445	protein C2656.10 l
46	0.7	849	2 B83349	probable ClpA/B-ty
47	0.7	1041	2 T31437	integrin alpha cha
48	0.7	1065	2 T25069	hypothetical prote
49	0.7	1146	2 S40311	integrin - fruit f
50	0.7	1153	1 RWHU1B	cell surface glyco
51	0.7	1163	1 RWHU1C	cell surface glyco
52	0.7	1179	2 A53213	integrin alpha-B c
53	0.7	1366	2 S57664	IgA-specific metal
54	0.7	1460	2 D81675	polymorphic membra
55	0.7	1664	2 S67250	DNA-directed RNA p
56	0.7	1802	2 H88444	protein C2656.12 l
57	0.7	2238	1 RRVUBY	genome polyprotein
58	0.6	47	2 D81686	hypothetical prote
59	0.6	74	2 T47376	hypothetical prote
60	0.6	89	2 I49515	B144 protein B - m
61	0.6	94	2 A99774	probable excisiona
62	0.6	99	2 F85636	probable excisiona
63	0.6	103	2 G75257	hypothetical prote
64	0.6	104	2 PH0141	integrin alpha cha
65	0.6	106	4 S57386	hypothetical prote
66	0.6	107	2 UC1127	hypothetical prote
67	0.6	109	2 C56413	major allergen cha
68	0.6	110	1 INRB1	insulin precursor
69	0.6	110	1 IPRT1	insulin 1 precursor
70	0.6	110	2 S30268	protein hdeA precu
71	0.6	110	2 F91177	protein hdeA precu
72	0.6	110	2 G86023	protein hdeA precu
73	0.6	114	2 C95870	hypothetical prote
74	0.6	115	2 D75434	hypothetical prote
75	0.6	132	2 B81215	conserved hypotet
76	0.6	133	2 H75386	hypothetical prote
77	0.6	139	1 C45051	lampin 2 precursor
78	0.6	139	2 A38612	insulin-like growt
79	0.6	141	1 HACH1	hemoglobin alpha-D
80	0.6	141	2 S56103	hemoglobin alpha-D
81	0.6	142	2 J70573	retinoic acid-indu
82	0.6	142	2 G87647	hypothetical prote
83	0.6	143	2 T05564	hypothetical prote
84	0.6	154	2 S14947	2S albumin - Brazi
85	0.6	154	2 F83074	conserved hypotet
86	0.6	158	2 G82494	conserved hypotet
87	0.6	160	1 F70351	NADH2 dehydrogenas
88	0.6	160	2 S58214	apoptosis inducer
89	0.6	161	2 D71903	hit family protein
90	0.6	163	2 S66795	probable membrane
91	0.6	163	2 B72225	conserved hypotet
92	0.6	171	1 B64498	conserved hypotet
93	0.6	176	2 AC0207	probable exported
94	0.6	176	2 T38925	cytochrome oxidase
95	0.6	176	2 T46709	hypothetical prote
96	0.6	176	2 AH1250	hypothetical prote
97	0.6	179	2 A64551	hypothetical prote
98	0.6	179	2 D71957	hypothetical prote
99	0.6	180	2 E70358	HupE hydrogenase r
100	0.6	182	2 E82967	hypothetical prote

ALIGNMENTS

```

RESULT 1
A:55348
Integrin alpha-1 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 15-Sep-2003
C:Accession: A55348
R:Kern, A.; Briesewitz, R.; Bank, I.; Marcantonio, E.E.
J. Biol. Chem. 269, 22811-22816, 1994
A:Title: The role of the I domain in ligand binding of the human integrin alpha-1beta-1.
A:Reference number: A55348; MUID:94357930; PMID:7521332
A:Accession: A55348
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-272 <K&R>
A:Cross-references: GB:U01114
F:155-230/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match      1.3%; Score 15; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 57 DIVIVLDGNSIYPW 71

RESULT 2
A:45226
Integrin alpha-1 chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
C:Accession: A45226
R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit.
A:Reference number: A45226; MUID:93155124; PMID:8428973
A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBI:P124326)
F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match      1.3%; Score 15; DB 2; Length 1151;
Best Local Similarity 100.0%; Pred. No. 6.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 144 DIVIVLDGNSIYPW 158

RESULT 3
A:35854
Integrin alpha-1 chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 15-Sep-2003
C:Accession: A35854; S11243
R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J.
J. Cell Biol. 111, 709-720, 1990
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
A:Reference number: A35854; MUID:90338125; PMID:2380249
A:Accession: A35854
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1180 <IGN>
A:Cross-references: GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

```

```

Query Match      1.3%; Score 15; DB 2; Length 1180;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 172 DIVIVLDGNSIYPW 186

RESULT 4
I:45914
Integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin
A:Reference number: A54402; MUID:94193647; PMID:7511592
A:Accession: I45914
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: GB:L25886; NID:G439695; PIDN:AA859255.1; PID:G439696
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match      0.9%; Score 11; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
DB 495 TDVLLVGAPMY 505

RESULT 5
I:33998
Integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C:Species: Homo sapiens (man)
C>Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAK>
A:Cross-references: GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R:Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIIb and GPIIb
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:I24121; NID:G400342; PIDN:AAA16619.2; PID:G458535
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics.
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:I28031; OMIM:192974
A:Map position: 5q11.1-5q11.2

```


C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;10-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 0.9%; Score 11; DB 2; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.072; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 495 TDVLLVGAPMY 505
|||||
Db 506 TDVLLVGAPMY 516
|||||

RESULT 6
I51524
integrin alpha 2 subunit - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 15-Sep-2003
C;Accession: I51524
R;Whittaker, C.A.; DeSimone, D.W.
Development 117, 1239-1249, 1993
A;Title: integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryo
A;Reference number: I51524; MUID:94008528; PMID:8404528
A;Accession: I51524
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-74 <WHI>
A;Cross-references: GB:L10186; NID:G214538; PIDN:AAA16246.1; PID:G214539

Query Match 0.8%; Score 9; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAP 503
|||||
Db 14 TDVLLVGAP 22
|||||

RESULT 7
AH3567
permease [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AH3567
R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Nazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AH3567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-315 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53707.1; PID:G17984630; GSPDB:GNC00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10465
A;Map position: II

Query Match 0.8%; Score 9; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
|||||
Db 283 LLLALLVLA 291
|||||

RESULT 8

T43407
3-isopropylmalate dehydrogenase (EC 1.1.1.85) [validated] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C;Accession: T43407; T39850; T39871
R;Kikuchi, Y.; Kikazawa, Y.; Shimatake, G.; Yamamoto, M.
Curr. Genet. 14, 375-379, 1988
A;Title: The primary structure of the leu+ gene of Schizosaccharomyces pombe.
A;Reference number: Z22491; MUID:89106270; PMID:3063400
A;Accession: T43407
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-371 <KIK>
A;Cross-references: EMBL:M36910; NID:G173411; PIDN:AAA35316.1; PID:G173412
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21885
A;Accession: T39850
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-371 <WOO>
A;Cross-references: EMBL:AL031174; PIDN:CAA20106.1; GSPDB:GN000067
A;Experimental source: strain 972h; cosmid c1A4
R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21887
A;Accession: T39871
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 152-371 <VOL>
A;Cross-references: EMBL:AL021746; PIDN:CAA16840.1; GSPDB:GN000067; SPDB:SPBCL1E8.07c
A;Experimental source: strain 972h; cosmid c1E8
C;Genetics:
A;Gene: leu1; SPDB:SPBCL1E8.07c; SPBCL1A4.02c
A;Map position: 2
C;Function: <IPD1>
A;Description: catalyzes the oxidation of 3-carboxy-2-hydroxy-4-methylpentanoate to 3-c
A;Pathway: leucine biosynthesis
C;Function: <IPD2>
A;Description: EC 1.1.1.85 [validated, MUID:89106270]
C;Superfamily: 3-isopropylmalate dehydrogenase
C;Keywords: oxidoreductase

Query Match 0.8%; Score 9; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 DGVLGAVG 386
|||||
Db 69 DGVLGAVG 77
|||||

RESULT 9
S60658
legumin - Gnetum gnemon
N;Alternate names: 11S globulin; seed storage protein
C;Species: Gnetum gnemon
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C;Accession: S60658
R;Braun, H.; Horstmann, C.; Baumlein, H.
submitted to the EMBL Data Library, August 1995
A;Description: Legumins of the Gnetatae: characterization and evolutionary relationship
A;Reference number: S60658
A;Accession: S60658
A;Molecule type: mRNA
A;Residues: 1-607 <BRA>
A;Cross-references: EMBL:Z50779; NID:G949870; PIDN:CAA90642.1; PID:G949871
C;Superfamily: glycinin
C;Keywords: seed; storage protein

Query Match 0.8%; Score 9; DB 2; Length 607;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1161
|||||||
Db 10 LLLALLLV 18

RESULT 10

S44142
VWA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
C:Accession: S44142
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VWA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <EDE>
A:Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 0.8%; Score 9; DB 2; Length 1178;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAP 503
|||||||
Db 503 TDVLLVGAP 511

RESULT 11

A60822
Cytochrome P450 PB-3a - rat (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999
C:Accession: A60822; I55191
R:Ameliaz, Z.; Narbonne, J.F.; Wolf, C.R.; Robertson, L.W.; Oesch, F.
Biochem. Pharmacol. 37: 3245-3249, 1988
A:Title: Effect of nutritional imbalances on cytochrome P-450 isozymes in rat liver.
A:Reference number: A60822; MUID:88293549; PMID:3041969
A:Accession: A60822
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <AME>
R:Hashimoto, T.; Matsumoto, T.; Nishizawa, M.; Kawabata, S.
J. Biochem. 103: 487-492, 1988
A:Title: A mutant rat strain deficient in induction of a phenobarbital-inducible form of
A:Reference number: I55191; MUID:88273074; PMID:2839467
A:Accession: I55191
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-20 <RES>
A:Cross-references: GB:D00250; NID:g20727; PIDN:BAA00181.1; PID:g220728
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
|||||||
Db 6 LLLALLLV 13

RESULT 12

A34259
Cytochrome P450mt4, phenobarbital-inducible, mitochondrial, hepatic - rat (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Mar-1999
C:Accession: A34259
R:Shayiq, R.M.; Avadhani, N.G.
Biochemistry 29: 866-873, 1990
A:Title: A phenobarbital-inducible hepatic mitochondrial cytochrome P-450 immunochemical
A:Reference number: A34259; MUID:90254127; PMID:2340279
A:Accession: A34259
A:Molecule type: protein
A:Residues: 1-42 <SHA>
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; heme; liver; mitochondrion; monooxygenase; oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
|||||||
Db 6 LLLALLLV 13

RESULT 13

I51527
Integrin alpha 5 subunit - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 29-Sep-1999
C:Accession: I51527
R:Whittaker, C.A.; Desimone, D.W.
Development 117: 1239-1249, 1993
A:Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryo
A:Reference number: I51524; MUID:94008528; PMID:8404528
A:Accession: I51527
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-76 <WHI>
A:Cross-references: GB:I10191; NID:g214544; PIDN:AAA16249.1; PID:g214545
C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 RFGSSIAS 548
|||||||
Db 63 RFGSSIAS 70

RESULT 14

A43537
Heat-stable antigen M1/69-J11d precursor - mouse
N:Alternate names: CD24 protein; nectadin
C:Species: Mus musculus (house mouse)
C:Date: 06-Nov-1992 #sequence_revision 06-Nov-1992 #text_change 31-Jan-2000
C:Accession: A43537; I48287; S15784; S15783; S43709; S32240; S33129
R:Kay, R.; Takei, F.; Humphries, R.K.
J. Immunol. 145: 1952-1959, 1990
A:Title: Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens.
A:Reference number: A43537; MUID:90361906; PMID:2118158
A:Accession: A43537
A:Molecule type: mRNA
A:Residues: 1-76 <KAY>
A:Cross-references: GB:M58661; NID:g198985; PIDN:AAA39481.1; PID:g198986
R:Wenger, R.H.; Rochelle, J.M.; Seldin, M.F.; Kohler, G.; Nielsen, P.J.
J. Biol. Chem. 268: 23345-23352, 1993
A:Title: The heat stable antigen (mouse CD24) gene is differentially regulated but has a
A:Reference number: A48876; MUID:94043127; PMID:8226859
A:Accession: I48287
A:Status: translation not shown; translated from GB/EMBL/DBJ

A:Residues: 1-76 <KAY>
A:Cross-references: GB:M58661; NID:g198985; PIDN:AAA39481.1; PID:g198986
R:Wenger, R.H.; Rochelle, J.M.; Seldin, M.F.; Kohler, G.; Nielsen, P.J.
J. Biol. Chem. 268: 23345-23352, 1993
A:Title: The heat stable antigen (mouse CD24) gene is differentially regulated but has a
A:Reference number: A48876; MUID:94043127; PMID:8226859
A:Accession: I48287
A:Status: translation not shown; translated from GB/EMBL/DBJ

RESULT 12

A34259
Cytochrome P450mt4, phenobarbital-inducible, mitochondrial, hepatic - rat (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)

A;Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A;Reference number: S15783; MUID:91209380; PMID:2019286
A;Accession: S15784
A;Molecule type: DNA
A;Residues: 1-76 <WE3>
A;Cross-references: EMBL:X56469; NID:G51439; PIDN:CAA39841.1; PID:G51440
A;Accession: S15783
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-76 <WE2>
A;Cross-references: EMBL:X53825
R;Nielsen, P.J.
submitted to the EMBL Data Library, July 1990
A;Reference number: S19111
A;Accession: S43709
A;Molecule type: mRNA
A;Residues: 1-76 <NTE>
A;Cross-references: EMBL:X53825; NID:G55441; PIDN:CAA37822.1; PID:G55442
C;Genetics:
A;Gene: CD24
A;Map position: 10
A;Introns: 23/3
C;Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosphatidyl
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-56/Product: heat-stable antigen M169-Jld #status predicted <MAT>
F;57-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;56/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 15
I53107
CD24 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
A;Accession: I53107; S25146
R;Shirasawa, T.; Akashi, T.; Sakamoto, K.; Takahashi, H.; Maruyama, N.; Hirokawa, K.
Dev. Dyn. 198, 1-13, 1993
A;Title: Gene expression of CD24 core peptide molecule in developing brain and developin
A;Reference number: I53107; MUID:94122434; PMID:8292828
A;Accession: I53107
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-76 <RES>
A;Cross-references: EMBL:Z11663; NID:G55901; PIDN:CAA7731.1; PID:G55902
C;Keywords: phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 16
A48996
B cell surface antigen CD24 precursor - human
N;Alternate names: cluster-w4 antigen; signal transducer CD24
C;Species: Homo sapiens (man)
C;Date: 19-Dec-1993 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
A;Accession: I56114; A48996; I54201; I72676
R;Kay, R.; Rosten, P.M.; Sexton, T.L.; Kay, R.; Humphries, R.K.
J. Immunol. 147, 1412-1416, 1991
A;Title: CD24, a signal transducer modulating B cell activation responses, is a very shd

A;Reference number: I56114; MUID:91332450; PMID:1831224
A;Accession: I56114
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-80 <RES>
A;Cross-references: GB:L33930; NID:G500848; PID:G500849; GB:M58664; NID:G180167; PID:G180167
R;Jackson, D.; Waibel, R.; Weber, E.; Bell, J.; Stahel, R.A.
Cancer Res. 52, 5264-5270, 1992
A;Title: CD24, a signal-transducing molecule expressed on human B cells, is a major sur
A;Reference number: A48996; MUID:93007871; PMID:1327504
A;Accession: A48996
A;Molecule type: mRNA
A;Residues: 1-56, 'V', 58-80 <JAC>
A;Cross-references: GB:X69397; GB:S44888; NID:G396167; PIDN:CAA49195.1; PID:G396168
A;Experimental source: small cell lung carcinoma line SW2
A;Note: sequence extracted from NCBI backbone (NCBIN:114635, NCBIP:114636)
A;Note: both 57-Val and 57-Ala were found in small cell carcinoma line DC571
R;Hough, M.R.; Rosten, P.M.; Sexton, T.L.; Kay, R.; Humphries, R.K.
Genomics 22, 154-161, 1994
A;Title: Mapping of CD24 and homologous sequences to multiple chromosomal loci.
A;Reference number: I54201; MUID:95048364; PMID:7959762
A;Accession: I54201
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 5-11, 'W', 13-43, 'T', 45-80 <RE2>
A;Cross-references: GB:S75311; NID:G933885; PIDN:AAD14170.1; PID:G4261870
C;Genetics:
A;Gene: GDB:CD24
A;Cross-references: GDB:383795; OMIM:600074
A;Map position: 6q21-6q21
C;Keywords: B-cell; blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinosit
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-67/Product: B cell surface antigen CD24 #status predicted <MAT>
F;68-80/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;67/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 17
AF0856
conserved hypothetical protein STY3056 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
A;Accession: AF0856
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0856
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06037.1; PID:GI6504004; GSPDB:GN00176
C;Genetics:
A;Gene: STY3056
C;Superfamily: hypothetical protein HI0673

Query Match 0.7%; Score 8; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLALL 1160

Db 6 LLLLLALLV 13
|||||

RESULT 18

B75301
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: B75301
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75301
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <WHI>
A:Cross-references: GB:AE002054; GB:AE000513; NID:g6460010; PIDN:AAF11763.1; PID:g646001
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2210
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR2210

Query Match 0.7%; Score 8; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
|||||
Db 7 GLLLLALL 14

RESULT 19

S15785
heat-stable antigen-related hypothetical protein HSA-C - mouse
C:Species: Mus musculus (house mouse)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S15785
R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.
Eur. J. Immunol. 21, 1039-1046, 1991
A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A:Reference number: S15783; MUID:91209380; PMID:2019286
A:Accession: S15785
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <HEN>
A:Cross-references: EMBL:X56486; NID:g51441; PIDN:CAA39843.1; PID:g51442
A:Note: the authors translated the codon TTC for residue 87 as Ser and AGA for residue 1
A:Note: the authors did not translate the codon for residue 101
C:Genetics:
A:Introns: #status absent
C:Keywords: glycoprotein; membrane protein; phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
|||||
Db 12 GLLLLALL 19

RESULT 20

E71228
hypothetical protein PH0087 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: E71228
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekim

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi.
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71228
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-224 <KAW>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29156.1; PID:g3256473
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0087

Query Match 0.7%; Score 8; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 512 RGVVYVE 519
|||||
Db 105 RGVVYVE 112

RESULT 21

S76800
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76800
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76800
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <KAN>
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAA18712.1; PID:g165380
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: conserved hypothetical protein MTH747

Query Match 0.7%; Score 8; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1073 RLVPNQEI 1080
|||||
Db 183 RLVPNQEI 190

RESULT 22

S15102
eosinophil major basic protein 2 precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: S15102; S18501
R:Aoki, I.; Shindoh, Y.; Nishida, T.; Nakai, S.; Hong, Y.M.; Mio, M.; Saito, T.; Tasaka,
PEBS Lett. 282, 56-60, 1991
A:Title: Comparison of the amino acid and nucleotide sequences between human and two gui
A:Reference number: S15102; MUID:91224343; PMID:2026266
A:Accession: S15102
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-234 <AOK>
A:Cross-references: DDBJ:D00817; NID:g3135095; PIDN:BAA00697.1; PID:g220293
A:Accession: S18501
A:Molecule type: protein

A:Residues: 116, 'X', 118-134, 'X', 136-137, 'X', 139-145; 161-176; 181-200 <AOK2>
C:Superfamily: eosinophil major basic protein precursor; C-type lectin homology
C:Keywords: antibiotic; chondroitin sulfate proteoglycan; cytotoxin; eosinophil; glycom

F;1-16/Domain: signal sequence #status predicted <SIG>
 F;17-115/Domain: propeptide #status predicted <PRO>
 F;116-234/Product: eosinophil major basic protein 2 #status experimental <MAT>
 F;117-232/Domain: C-type lectin homology <LCH>
 F;24-25/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F;59/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
 F;135-232,209-224/disulfide bonds: #status predicted

Query Match 0.7%; Score 8; DB 1; Length 234;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 |||||
 Db 4 LLLALLV 11

RESULT 23
 S22090
 catechol O-methyltransferase (EC 2.1.1.6) - rat
 N/Alternate names: catechol-O-methyltransferase
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
 C/Accession: S22090; JQ0787; I52356
 R/Tenhuunen, J.; Salminen, M.; Jalanko, A.; Ukkonen, S.; Ulmanen, I.
 submitted to the EMBL Data Library, June 1992
 A/Description: Structure of the rat catechol-O-methyltransferase gene: Separate promoter
 A/Reference number: S22090
 A/Accession: S22090
 A/Molecule type: DNA
 A/Residues: 1-264 <RES>
 A/Cross-references: EMBL:Z12651; NID:G55889; PIDN:CAA78276.1; PID:G55890
 R/Salminen, M.; Lundstroem, K.; Tilgmann, C.; Savolainen, R.; Kalkkainen, N.; Ulmanen, I.
 Genes 93, 241-247 1990
 A/Title: Molecular cloning and characterization of rat liver catechol-O-methyltransferase
 A/Reference number: JQ0787; MUID:91033034; PMID:2227437
 A/Accession: JQ0787
 A/Molecule type: mRNA
 A/Residues: 44-264 <SAL>
 A/Cross-references: GB:M50753; NID:G203336; PIDN:AAA40881.1; PID:G203337
 A/Experimental source: liver
 R/Tenhuunen, J.; Ulmanen, I.
 Biochem. J. 296, 595-600, 1993
 A/Title: Production of rat soluble and membrane-bound catechol O-methyltransferase forms
 A/Reference number: I52356; MUID:94107221; PMID:8280056
 A/Accession: I52356
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-264 <RES>
 A/Cross-references: EMBL:Z12651; NID:G55889; PIDN:CAA78276.1; PID:G55890
 C/Comment: This enzyme catalyzes the transfer of a methyl group from S-adenosyl-L-methionine to catechol
 A/Gene: COMT
 C/Superfamily: caffeoyl-CoA 3-O-methyltransferase
 C/Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 264;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALLL 1159
 |||||
 Db 9 GLLALLL 16

RESULT 24
 AG2637
 conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C58)
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C/Accession: AG2637
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AG2637
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-308 <KUR>
 A/Cross-references: GB:AE008698; PIDN:AAL41517.1; PID:G17738846; GSPDB:GN00186
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: Atu0498
 A/Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
 |||||
 Db 275 LLALLVLA 282

RESULT 25
 F97419
 BH1459 conserved hypothetical protein (AF001512) [imported] - Agrobacterium tumefaciens
 C/Species: Agrobacterium tumefaciens
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C/Accession: F97419
 A/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
 Science 294, 2323-2328, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
 A/Reference number: A97359; MUID:21608551; PMID:11743194
 A/Accession: F97419
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-356 <KUR>
 A/Cross-references: GB:AE007869; PIDN:AAK86311.1; PID:G15155427; GSPDB:GN00169
 C/Genetics:
 A/Gene: AGR_C 881
 A/Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
 |||||
 Db 323 LLALLVLA 330

RESULT 26
 A85016
 hypothetical protein AT4G01220 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C/Accession: A85016
 R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr
 Nature 402, 769-777, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001; MUID:20083488; PMID:10617198
 A/Accession: A85016
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-360 <STO>
 A/Cross-references: GB:NC_001268; NID:G7267619; PIDN:CAB80931.1; GSPDB:GN00140

QY 1152 GLLALLL 1159
 |||||
 Db 9 GLLALLL 16

RESULT 24
 AG2637
 conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C58)
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C/Accession: AG2637
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

Query Match 0.7%; Score 8; DB 2; Length 415;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157
 |||||
 Db 388 LGGLLLLA 395

RESULT 31
 AG3504
 C:Species: dihydrofolate synthase (EC 6.3.2.12) [imported] - Brucella melitensis (strain 16M)
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
 C:Accession: AG3504
 R:DelVecchio, V.G.; Kapatral, R.J.; Redkar, R.J.; Patra, G.; Mujar, C.; Ios, T.; Ivanova,
 ; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756588
 A:Accession: AG3504
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-442 <KOR>
 A:Cross-references: GB:AB008917; PIDN:AAL53202.1; PID:gl7984076; GSPDB:GNO0190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMBI2021
 A:Map position: I
 C:Superfamily: folsypolyglutamate synthase
 C:Keywords: ligase

Query Match 0.7%; Score 8; DB 2; Length 442;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 NEDGLIDL 622
 |||||
 Db 241 NEDGLIDL 248

RESULT 32
 T01721
 C:Species: hypothetical protein A_IC002N01.8 - Arabidopsis thaliana
 C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Mar-1999
 C:Accession: T01721
 R:Scheet, P.; Maggi, L.
 submitted to the EMBL Data Library, June 1997
 A:Description: The sequence of A. thaliana IC002N01.
 A:Reference number: Z14407
 A:Accession: T01721
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-444 <SCH>
 A:Cross-references: EMBL:AF007269; NID:g2191126; PID:g2191131
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 68/1; 235/3; 302/3; 374/3
 A:Note: A_IC002N01.8

Query Match 0.7%; Score 8; DB 2; Length 444;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLALLV 1160
 |||||
 Db 115 LLLLALLV 122

RESULT 33
 T47107

benzaldehyde dehydrogenase (NAD) (EC 1.2.1.28) [validated] - Pseudomonas putida plasmid
 C:Species: Pseudomonas putida
 C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 03-Jun-2002
 C:Accession: T47107; S13388
 R:Inoue, J.; Shaw, J.P.; Bekik, M.; Harayama, S.
 J. Bacteriol. 177, 1196-1201, 1995
 A:Title: Overlapping substrate specificity of benzaldehyde dehydrogenase (the xylC gene
 Pseudomonas putida.
 A:Reference number: Z24352; MUID:95173094; PMID:7868591
 A:Accession: T47107
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-487 <INO>
 A:Cross-references: EMBL:U15151; NID:g555994; PIDN:AAA66218.1; PID:g555995
 R:Chalmers, R.M.; Keen, J.N.; Fewson, C.A.
 Biochem. J. 273, 99-107, 1991
 A:Title: Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenases fr
 uene pathway in Pseudomonas putida.
 A:Reference number: S13298; MUID:91113163; PMID:1989592
 A:Accession: S13388
 A:Molecule type: protein
 A:Residues: 1-44 <CHA>
 A:Experimental source: strain MT53
 C:Genetics:
 A:Gene: xylC
 A:Genome: plasmid
 A:Note: TOL plasmid pWMO
 C:Function:
 A:Description: EC 1.2.1.28 [validated, MUID:95173094]
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 C:Keywords: oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 487;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 ALGNVAIL 633
 |||||
 Db 171 ALGNVAIL 178

RESULT 34
 O4RTPB
 C:Species: cytochrome P450 2B1 - rat
 N:Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible
 N:Contains: unspecific monooxygenase (EC 1.14.14.1)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 18-Aug-1982 #sequence_revision 17-May-1996 #text_change 03-Mar-2000
 C:Accession: A00176; A54251; A22363; A29288; S03854; A92255; I54796
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
 A:Title: Primary structure of a cytochrome p-450: coding nucleotide sequence of phenoba
 A:Reference number: A93912; MUID:82222224; PMID:6953431
 A:Accession: A00176
 A:Molecule type: mRNA
 A:Residues: 6-491 <PUJ>
 A:Cross-references: EMBL:J00719; NID:g203752; PIDN:AAA1024.1; PID:g203753
 A:Note: the authors translated the codon GAT for residue 166 as Glu, CTG for residue 29
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
 A:Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobar
 A:Reference number: A93925
 A:Contents: annotation
 A:Note: the mistranslations shown in reference A93912 are acknowledged
 R:Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg, P
 Biochemistry 33, 3766-3771, 1994
 A:Title: Identification of active-site peptides from (3)H-labeled 2-ethynylanthralene.
 A:Reference number: A54251; MUID:94190899; PMID:8142377
 A:Accession: A54251
 A:Molecule type: protein
 A:Residues: 290-301, 'X', <ROB>
 R:Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.
 J. Biol. Chem. 260, 7980-7984, 1985

A>Title: Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in rat
 A:Reference number: A22363; MUID:85234490; PMID:2995270
 A:Accession: A22363
 A:Molecule type: DNA
 A:Residues: 1-91, 'P', 93-204, 'R', 206-327, 'V', 329-356, 'H', 358-391, 'R', 393-415, 'V', 417-433, 'A';
 A:Cross-references: GB:L00320; NID:G203816; PIDN:AAA41046.1; PID:G203818
 A:Note: the authors translated the codon CAG for residue 57 as Gly, CTT for residue 92 as Arg, AAA for residue 236 as Leu, AGC for residue 259 as Asn, GTT for residue 328 as T
 R:Angarajan, P.N.; Ravishanker, H.; Padmanaban, G.
 Biochem. Biophys. Res. Commun. 144, 258-263, 1987
 A>Title: Isolation of a cytochrome P-450e gene variant and characterization of its 5' flanking region
 A:Reference number: A29298; MUID:87213174; PMID:3579906
 A:Accession: A29298
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-57 <RAN>
 R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
 Arch. Biochem. Biophys. 270, 23-32, 1989
 A>Title: Antibodies targeted against hypervariable and constant regions of cytochromes P-450
 A:Reference number: S03854; MUID:89192373; PMID:2539047
 A:Accession: S03854
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-18; 146-160, 'E', 162-165; 166, 330-361; 362-380; 402-423 <OES>
 R:Botelho, L.H.; Ryan, D.E.; Levin, W.
 J. Biol. Chem. 254, 5635-5640, 1979
 A>Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene
 A:Reference number: A92255; MUID:79194111; PMID:109438
 A:Accession: A92255
 A:Molecule type: protein
 A:Residues: 1-3, 'T', 5-22 <BOT>
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Taniguchi, T.; Muramatsu, M.
 Int. Symp. Princess Takamatsu Cancer Res. Fund 12, 31-40, 1982
 A>Title: Molecular cloning and coding nucleotide sequence of complementary DNA of cytochrome P-450
 A:Reference number: 154796; MUID:83160754; PMID:6300027
 A:Accession: 154796
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 6-491 <RES>
 A:Cross-references: GB:M37134; NID:G203784; PIDN:AAC42028.1; PID:G203785
 C:Genetics:
 A:Gene: CYP2B1
 A:Introns: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metalloprotein
 F/295-458/Domain: cytochrome P450 homology <P45>
 F/302/Active site: Thr #status Predicted
 F/436/Binding site: heme iron (Cys) (axial ligand) #status predicted
 Query Match 0.7%; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALLLV 1160
 Db 6 LLLALLLV 13
 RESULT 35
 O4RTP2
 N:Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cytochrome P450 2B2 - rat
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000
 C:Accession: A21162; A00177; B00176; B92255; S15589; A21872; A32736; S03855; I59060
 R:Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983
 A>Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.
 A:Reference number: A21162; MUID:83247397; PMID:6306654
 A:Accession: A21162

A:Molecule type: DNA
 A:Residues: 1-472, 'M', 474-491 <MIZ>
 A:Cross-references: EMBL:J00728; NID:G203845; PIDN:AAA41056.1; PID:G203847
 A:Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue 3 as Arg
 R:Prev, A.B.; Waxman, D.J.; Kreibich, G.
 J. Biol. Chem. 260, 15253-15265, 1985
 A>Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme P-450
 A:Reference number: A00177; MUID:86059379; PMID:3877725
 A:Accession: A00177
 A:Molecule type: protein
 A:Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <PRE>
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
 A>Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phenobarbital
 A:Reference number: A93912; MUID:82222224; PMID:6953431
 A:Accession: B00176
 A:Molecule type: mRNA
 A:Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477, 'G'
 A:Note: nucleotide sequence for residues 1-5 is not given
 A:Note: the authors translated the codon GAT for residue 166 as Glu, CTG for residue 292 as Arg
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
 A>Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobarbital
 A:Reference number: A93925
 A:Contents: annotation; revisions
 A:Note: the mistranslations in reference A93912 are acknowledged
 R:Botelho, L.H.; Ryan, D.E.; Levin, W.
 J. Biol. Chem. 254, 5635-5640, 1979
 A>Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene
 A:Reference number: A92255; MUID:79194111; PMID:109438
 A:Accession: B92255
 A:Molecule type: protein
 A:Residues: 1-3, 'T', 5-22 <BOT>
 R:LaCroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.
 Gene 86, 201-207, 1990
 A>Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB21).
 A:Reference number: S15589; MUID:90215299; PMID:2323573
 A:Accession: S15589
 A:Molecule type: mRNA
 A:Residues: 105-113, 'P', 115-274, 'VSPAWRE', 275-321, 'E', 323-491 <LAC>
 A:Cross-references: EMBL:M34452; NID:G203679; PIDN:AAA41004.1; PID:G203680
 A:Note: translation of the nucleotide sequence is not complete
 R:Phillips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.
 Gene 24, 41-52, 1983
 A:Reference number: A21872
 A:Accession: A21872
 A:Molecule type: mRNA
 A:Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>
 R:Atfalter, M.; Anderson, A.
 Biochem. Biophys. Res. Commun. 118, 655-662, 1984
 A>Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450
 A:Reference number: A32736; MUID:84153837; PMID:6322758
 A:Accession: A32736
 A:Molecule type: mRNA
 A:Residues: 385-491 <AFF>
 A:Cross-references: GB:K01626; NID:G203782; PIDN:AAA41037.1; PID:G203783
 R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
 Arch. Biochem. Biophys. 270, 23-32, 1989
 A>Title: Antibodies targeted against hypervariable and constant regions of cytochromes P-450
 A:Reference number: S03854; MUID:89192373; PMID:2539047
 A:Accession: S03855
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 329-358, 'AS', 361; 362; 363-380; 402-423 <OES>
 R:Archison, M.L.; Adesnik, M.
 Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986
 A>Title: Gene conversion in a cytochrome P-450 gene family.
 A:Reference number: I59060; MUID:86205943; PMID:3458196
 A:Accession: I59060
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 323-431 <RES>

Science 294, 849-852, 2001
A;Authors: Krefzt, J.; Kuhn, M.; Kuner, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; UID:21537279; PMID:11679669
A;Accession: AC1314
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-547 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99993.1; PID:g16411368; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1915
C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;
Best Local Similarity 100.0%; Pred.No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 MIVITDGE 273
|||||
DB 141 MIVITDGE 148

RESULT 38
AC1686
malolactic enzyme (malate dehydrogenase) homolog lin2029 [imported] - Listeria innocua
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AC1686
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. Science 294, 849-852, 2001
A;Authors: Krefzt, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; UID:21537279; PMID:11679669
A;Accession: AC1686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-547 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97259.1; PID:g16414530; GSPDB:GN00178
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin2029
C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;
Best Local Similarity 100.0%; Pred.No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 MIVITDGE 273
|||||
DB 141 MIVITDGE 148

RESULT 39
MWLEB
E1 protein - bovine papillomavirus type 1
C;Species: bovine papillomavirus type 1
C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 28-Jul-2000
C;Accession: A03663
R;Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H. Nature 299, 529-534, 1992
A;Title: The primary structure and genetic organization of the bovine papillomavirus ty A;Reference number: A93289; UID:83012974; PMID:6289124
A;Accession: A03663
A;Molecule type: DNA
A;Residues: 1-605 <CHE>
A;Cross-references: GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:g60965; PIDN:CAB465 C;Superfamily: Papillomavirus E1 protein
C;Keywords: early protein

```

Query Match
Best Local Similarity 0.7%; Score 8; DB 1; Length 605;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
DB 337 LAAGSDSN 344
|||||

RESULT 40
A72429
oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: A72429
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; PMID:99287316; PMID:10360571
A:Accession: A72429
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-606 <ARN>
A:Cross-references: GB:AE001690; GB:AE000512; NID:G4980496; PIDN:AA035125.1; PID:G498051
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0031
C:Superfamily: dipeptide transport protein

Query Match
Best Local Similarity 0.7%; Score 8; DB 2; Length 606;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 9 LLLALLVL 16
|||||

RESULT 41
A69845
Nat/H+ antiporter homolog yjBQ - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: A69845
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Better
C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seroi
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; PMID:98044033; PMID:9384377
A:Accession: A69845
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-614 <XUN>
A:Cross-references: GB:Z99110; GB:AL009126; NID:G2633472; PIDN:CA013021.1; PID:e1183184;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjBQ

Query Match
0.7%; Score 8; DB 2; Length 614;

```

```

Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 313 LLLALLLV 320
|||||

RESULT 42
WLWLS2
E1 protein - bovine papillomavirus type 2
C:Species: bovine papillomavirus type 2
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 12-Jun-1998
C:Accession: C31169
R:Groff, D.E.; Mitra, R.; Lancaster, W.D.
submitted to GenBank, May 1988
A:Reference number: A94519
A:Accession: C31169
A:Molecule type: DNA
A:Residues: 1-620 <GRO>
A:Cross-references: GB:M20219; GB:M19551; NID:G332996
C:Superfamily: papillomavirus E1 protein
C:Keywords: early protein; glycoprotein
F172109;173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 620;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
DB 352 LAAGSDSN 359
|||||

RESULT 43
AC0527
ferrichrome transport protein PhuB precursor STY0221 [imported] - Salmonella enterica su
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0527
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:21534947; PMID:11677608
A:Accession: AC0527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-685 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01355.1; PID:gl6501482; GSPDB:GN00176
C:Genetics:
A:Gene: STY0221
C:Superfamily: vitamin B12 transport protein bruc

Query Match
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 36 LLLALLLV 43
|||||

RESULT 44
T16306
hypothetical protein F40F4.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16306
R:Wilson, R.

```

```
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F40F4.
A:Reference number: Z18493
A:Accession: T16306
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-697 <WIL>
A:Cross-references: EMBL:U40420; NID:g1065513; PID:g1065515; PIDN:AAA81431.1; CESP:F40F4
C:Genetics:
A:Introns: 14/2; 86/3; 135/3; 223/1; 244/1; 331/1; 365/1; 396/1; 435/1; 559/3; 627/2; 69
Query Match 0.7%; Score 8; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 YNDVVVGA 564
DB 593 YNDVVVGA 600
RESULT 45
A88445
protein C26E6.10 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88445
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-739 <STC>
A:Cross-references: GB:chr_III; PIDN:AAA21163.1; PID:g532810; GSPDB:GN00021; CESP:C26E6.
C:Genetics:
A:Gene: C26E6.10
A:Map position: 3
Query Match 0.7%; Score 8; DB 2; Length 739;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 430 TSVVSSRQ 437
DB 234 TSVVSSRQ 241
RESULT 46
E83349
probable C1pA/B-type proteinase PA2371 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83349
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:2043737; PMID:10984043
A:Accession: E83349
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-849 <STO>
A:Cross-references: GB:AE004663; GB:AE004091; NID:g9948405; PIDN:AG05759.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2371
C:Superfamily: endopeptidase Clp ATP-binding chain
Query Match 0.7%; Score 8; DB 2; Length 849;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1152 GLLLLALL 1159
DB 111 GLLLLALL 118
RESULT 47
TJ1437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: TJ1437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: TJ1437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC3572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain
Query Match 0.7%; Score 8; DB 2; Length 1041;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 498 LLVGAPMY 505
DB 308 LLVGAPMY 315
RESULT 48
T25068
hypothetical protein T21C9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25068
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19977
A:Accession: T25068
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1065 <WIL>
A:Cross-references: EMBL:Z73098; PIDN:CAA97331.1; GSPDB:GN00023; CESP:T21C9.2
A:Experimental source: clone T21C9
C:Genetics:
A:Gene: CESP:T21C9.2
A:Map position: 5
A:Introns: 67/2; 147/2; 207/3; 242/3; 293/3; 825/2; 897/1; 1032/2
Query Match 0.7%; Score 8; DB 2; Length 1065;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 866 FASLIQKE 873
DB 457 FASLIQKE 464
RESULT 49
S40311
integrin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
C:Accession: S40311
```

R;Wehrli, M.; DiAntonio, A.; Fearnley, I.M.; Smith, R.J.; Wilcox, M. Mech. Dev. 43, 21-36, 1993

A>Title: Cloning and characterization of alpha(P51), a novel *Drosophila melanogaster* integrin alpha chain; neutrophil adherence receptor alpha chain

A:Reference number: S40311; MUID:94059764; PMID:8240969

A:Accession: S40311

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1146 <WEH>

A:CROSS-references: ENBL:X73975; NID:9440143; PIDN:CA52155.1; PID:9440144

A:Gene: FlyBase:FBgn0004456

C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1146;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLLL 1158
|||||
DB 1094 GGLLLLLL 1101

RESULT 50

RWHUIC

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1; integrin alpha chain; neutrophil adherence receptor alpha chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

A:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; 152567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A. J. Biol. Chem. 263, 12403-12411, 1988

A>Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b)

A:Reference number: A31108; MUID:98315033; PMID:2457594

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:CROSS-references: GB:J03925; NID:9187284; PIDN:AAA59544.1; PID:9307148

A>Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G. J. Cell Biol. 106, 2153-2158, 1988

A>Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mac-1

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-965, P', 967-1153 <ARN>

A:CROSS-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:9186935; PIDN:AAA594

A>Note: the authors translated the codon TAC for residue 1129 as Thr

R:Shelley, C.S.; Arnaout, M.A. Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A>Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:CROSS-references: GB:M76724; NID:9180018; PIDN:AAA58410.1; PID:9553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G. Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A>Title: Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion receptor

A:Reference number: A94193; MUID:88130151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:CROSS-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J. Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A>Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:CROSS-references: GB:J04145; NID:9189068; PIDN:AAA59903.1; PID:9386975

A>Note: part of this sequence was confirmed by protein sequencing

R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G. J. Immunol. 150, 480-490, 1993

A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha chain during evolution

A:Reference number: A46526; MUID:93123748; PMID:8419480

A:Accession: A46526

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-499, 501-1153 <FLE>

A:CROSS-references: GB:S52227; NID:9263047; PIDN:AA24821.1; PID:9263049

A>Note: the last three bases of intron 13, CAG, are included in some but not all mature cDNAs

R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A. Biochim. Biophys. Acta 874, 368-371, 1986

A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species

A:Reference number: A30664; MUID:87076671; PMID:3539202

A:Accession: A30691

A:Molecule type: protein

A:Residues: 17-31 <PIE>

A:Experimental source: granulocytes

R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G. Blood 79, 865-870, 1992

A>Title: Characterization of the myeloid-specific CD11b promoter.

A:Reference number: 152567; MUID:92144986; PMID:1346576

A:Accession: 152567

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-9 <RES>

A:CROSS-references: GB:M84477; NID:9180184; PIDN:AAA51960.1; PID:9553219

C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1

C:Genetics:

A:Gene: GDB:ITGAM; CR3A

A:CROSS-references: GDB:120599; OMIM:120980

A:Map position: 16p11.2-16p11.2

A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homolog

C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <VAT>

F:17-1108/Domain: extracellular #status predicted <EXT>

F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>

F:465-473/Region: calcium/magnesium binding #status predicted

F:530-538/Region: calcium/magnesium binding #status predicted

F:593-601/Region: calcium/magnesium binding #status predicted

F:1109-1134/Domain: transmembrane #status predicted <TM>

F:1135-1153/Domain: intracellular #status predicted <INT>

F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 0.7%; Score 8; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLLL 1158
|||||
DB 1116 GGLLLLLL 1123

RESULT 51

RWHUIC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 22-Jun-1999

A:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A. J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A;Residues: 1-1163 <COR>
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36594
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166545; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755 'L', 757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <EXT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carboxydrate (Asn) (covalent) #status pre

Query Match 0.7%; Score 8; DB 1; Length 1163;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1151 GGLLLAL 1158
|||||
Db 1115 GGLLLAL 1122
|||||

RESULT 52
A53213
Integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6015-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: GB:L25851; NID:G457244; PID:G457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;195-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 0.7%; Score 8; DB 2; Length 1179;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 380 VLLGAVGA 387
|||||
Db 416 VLLGAVGA 423
|||||

RESULT 53
S57664
IGA-specific metalloendopeptidase homolog sepA precursor - Shigella flexneri
C;Species: Shigella flexneri
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Dec-2000
C;Accession: S57664; S69769; S69768
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.
submitted to the EMBL Data Library, February 1995
A;Description: Characterization of SepA, the major extracellular protein of Shigella fl
A;Reference number: S57664
A;Accession: S57664
A;Molecule type: DNA
A;Residues: 1-1366 <BEN>
A;Cross-references: EMBL:Z48219; NID:G886952; PIDN:CAA88252.1; PID:G886953
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.
Mol. Microbiol. 17, 123-135, 1995
A;Title: SepA, the major extracellular protein of Shigella flexneri: autonomous secreti
A;Reference number: S69768; MUID:96020667; PMID:7476198
A;Accession: S69769
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-66;1077-1366 <BEW>
A;Cross-references: EMBL:Z48219; NID:G886952; PIDN:CAA88252.1; PID:G886953
A;Experimental source: strain N90T (serotype 5)
A;Accession: S69768
A;Molecule type: protein
A;Residues: 57-61, Q', 63-72, 'X', 544-557, 'X', 1057-1068 <BEF>
A;Note: 6-Glu, 14-Thr, 18-Gly and 26-Gln were also found
C;Genetics:
A;Gene: sepA
C;Superfamily: IGA-specific metalloendopeptidase
C;Keywords: extracellular protein
F;1-56/Domain: signal sequence #status predicted <SIG>
F;57-1366/Product: IGA-specific metalloendopeptidase homolog sepA #status predicted <WA

Query Match 0.7%; Score 8; DB 2; Length 1366;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 100 KDNMRLGL 107
|||||
Db 1335 KDNMRLGL 1342
|||||

RESULT 54
DB1675
polymorphic membrane protein B/C family TC0695 [imported] - Chlamydia muridarum (strain
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: DB1675
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.P.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: DB1675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1460 <TET>
A;Cross-references: GB:AE002338; GB:AE002160; NID:G7190724; PIDN:AAF39511.1; PID:G71907
A;Experimental source: strain N-gg (MoPn)
C;Genetics:
A;Gene: TC0695

Query Match 0.7%; Score 8; DB 2; Length 1460;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 583 SILKTPKQ 590
|||||
Db 1311 SILKTPKQ 1318
|||||

RESULT 55

S67250
DNA-directed RNA polymerase (EC 2.7.7.6) I 190K chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: DNA-directed RNA polymerase A 190K chain; protein O6276; protein YOR3
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67250; A29926; S67396
R:Goffeau, A.; Purnelle, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67246

A:Accession: S67250

A:Molecule type: DNA

A:Residues: 1-1664 <GOF>

A:Cross-references: EMBL:Z75249; NID:gl420740; PIDN:CAA93665.1; PID:gl420741; MIPS:YOR34

A:Experimental source: strain S288C

R:Memet, S.; Gouy, M.; March, C.; Sentenac, A.; Buhler, J.M.

J. Biol. Chem. 263, 2830-2839, 1988

A:Title: RPA190, the gene coding for the largest subunit of yeast RNA polymerase A.

A:Reference number: A29926; MUID:88139333; PMID:2830265

A:Accession: A29926

A:Molecule type: DNA

A:Residues: 1-157; T, 159-1664 <MEM>

A:Cross-references: EMBL:J03530; NID:gl72214; PIDN:AAA34890.1; PID:gl72215

R:Purnelle, B.; Goffeau, A.

submitted to the EMBL Data Library, February 1996

A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast
ome I genes.

A:Reference number: S67392

A:Accession: S67396

A:Molecule type: DNA

A:Residues: 1-1664 <PUR>

A:Cross-references: EMBL:X95720; NID:gl199839; PIDN:CAA65029.1; PID:gl199844

C:Genetics:

A:Gene: SGD:RPA190

A:Cross-references: SGD:S0005868; MIPS:YOR341W

A:Map position: 15R

C:Superfamily: Trypanosoma DNA-directed RNA polymerase I largest chain

C:Keywords: DNA binding; nucleotidyltransferase; transcription; zinc finger

F:62-78/Region: zinc finger CCHH motif

Query Match

Best Local Similarity 100.0%; Pred. No. 1e+02; Length 1664;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

481 GSEITSVD 488

Db

8 GSEITSVD 15

RESULT 56

H88444

protein C26E6.12 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: H88444

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A:Reference number: A75000; MUID:99059613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: H88444

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1802 <STO>

A:Cross-references: GB:chr_III; PIDN:AAA21166.1; PID:G532813; GSPDB:GN00021; CESP:C26E6.

C:Genetics:

A:Gene: C26E6.12

A:Map position: 3

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e+02; Length 1802;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 430 TSVSSRRQ 437

Db 1297 TSVSSRRQ 1304

RESULT 57

RVUBY

genome polyprotein - bunyamwera virus

N:Alternate names: L protein

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: bunyamwera virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 11-Jun-1999

C:Accession: A33744

R:Elliot, R.M.

Virology 173, 426-436, 1989

A:Title: Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera

A:Reference number: A33744; MUID:90085791; PMID:2596023

A:Accession: A33744

A:Molecule type: genomic RNA

A:Residues: 1-2238 <ELL>

A:Cross-references: GB:X14383; NID:G58712; PIDN:CAA32553.1; PID:G58713

C:Genetics:

A:Map position: segment L

C:Superfamily: bunyavirus RNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase

Query Match

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Length 2238;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 FLINILKK 191

Db 1333 FLINILKK 1340

RESULT 58

DS1666

hypothetical protein TC0766 [imported] - Chlamydia muridarum (strain Nigg)

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000

C:Accession: DS1666

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: DS1666

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-47 <TET>

A:Cross-references: GB:AE002345; GB:AE002160; NID:G7190791; PIDN:AAF39569.1; PID:G719079

A:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0766

Query Match

Best Local Similarity 100.0%; Pred. No. 46; Length 47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALL 1159

Db 4 LLLALL 10

RESULT 59

T47376

hypothetical protein T5C2.40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47376

R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M

submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24463
A;Accession: T47376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <OBE>
A;Cross-references: EMBL:AL138664
A;Experimental source: cultivar Columbia; BAC clone T5C2
C;Genetics:
A;Map position: 3
A;Introns: 36/3; 55/3
A;Note: T5C2.40

Query Match 0.6%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 HCVFPLV 792
|||
DB 13 HCVFPLV 19

RESULT 60

I49515
B144 protein B - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49515
R;Teuge, I.; Shen, F.
Immunogenetics 26, 378-380, 1987
A;Title: A gene in the H-2S; H-2D interval of the major histocompatibility complex which
A;Reference number: I49514; MUID:88031493; PMID:3117682
A;Accession: I49515
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-89 <REG>
A;Cross-references: GB:M18187; NID:gl92097; PIDN:AAA37273.1; PID:gl92099

Query Match 0.6%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGJLLL 1156
|||
DB 36 LGGJLLL 42

RESULT 61

A99774
probable excisionase [imported] - Escherichia coli (strain O157:H7, substrain RMD 05099)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: A99774
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99774
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-94 <RAY>
A;Cross-references: GB:BA000007; PIDN:BA334584.1; PID:gl3360621; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECS1161

Query Match 0.6%; Score 7; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 637 PVQINA 643
|||

DB 18 PVQINA 24

RESULT 62

F85636
probable excisionase for bacteriophage BP-933W xisW [imported] - Escherichia coli (stra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <STO>
A;Cross-references: GB:AE005174; NID:gl2514277; PIDN:AAG55554.1; GSPDB:GN00145; UWGP:Z1
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: xisW

Query Match 0.6%; Score 7; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 637 PVQINA 643
|||
DB 23 PVQINA 29

RESULT 63

G75257
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: G75257
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <WHI>
A;Cross-references: GB:AE002086; GB:AE000513; NID:G6460395; PIDN:AAF2117.1; PID:G64604
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2573
A;Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163
|||
DB 15 ALLVLAL 21

RESULT 64

PH0141
integrin alpha chain - chicken (fragments)
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2003
C;Accession: PH0141
R;Syfrig, J.; Mann, K.; Paulsson, M.
Exp. Cell Res. 194, 165-173, 1991
A;Title: An abundant chick gizzard integrin is the avian alpha1-beta1 integrin heterodi
A;Reference number: PH0141; MUID:91224161; PMID:1851093

A:Accession: PH0141
A:Molecule type: protein
A:Residues: 1-104 <SYF>
A:Experimental source: gizzard
C:Comment: Integrins are a family of cellular adhesion receptors that mediate interaction
C:Keywords: cell adhesion

Query Match 0.6%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 QIGSYFG 481
DB 45 QIGSYFG 51

RESULT 65
S57386
hypothetical protein orf 00954 - Yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1995 #sequence_revision 05-Sep-1996 #text_change 20-Oct-2000
C:Accession: S57386
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than 11,975-986, 1995
A:Reference number: S57374; MUID:96021609; PMID:8533473
A:Accession: S57386
A:Status: nucleic acid sequence not shown; conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-106 <ZUM>
A:Cross-references: EMBL:X83121; NID:G600461; PIDN:CAA58196.1; PID:G600475
C:Comment: There is no evidence that this sequence is expressed.
C:Genetics:
A:Map position: 15L
C:Keywords: pseudogene

Query Match 0.6%; Score 7; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 SFLACSP 123
DB 17 SFLACSP 23

RESULT 66
JC1127
major allergen chain 2 precursor, short form - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: JC1127
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.
A:Title: Expression and genomic structure of the genes encoding FdI, the major allergen
A:Reference number: JC1126; MUID:92241678; PMID:1572548
A:Accession: JC1127
A:Molecule type: DNA
A:Residues: 1-107 <GRI>
A:Cross-references: GB:X62478; NID:G395406; PIDN:CAA44345.1; PID:G395407
A:Experimental source: skin
C:Genetics:
A:Gene: Ch2
A:Introns: 21/1; 81/3
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.6%; Score 7; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163

DB 4 ALLVLAL 10
RESULT 67
CS6413
major allergen Fel dI chain 2 precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 01-Dec-2000
C:Accession: CS6413; JC1145
R:Morgenstern, J.P.; Griffith, I.J.; Brauer, A.W.; Rogers, B.L.; Bond, J.F.; Chapman, M.
A:Title: Amino acid sequence of Fel dI, the major allergen of the domestic cat: protein
A:Reference number: A56413; MUID:92052157; PMID:1946388
A:Accession: CS6413
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-109 <MOR>
A:Cross-references: GB:M77341; NID:G163822; PIDN:AAC41616.1; PID:G163823
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.
A:Title: Expression and genomic structure of the genes encoding FdI, the major allergen
A:Reference number: JC1126; MUID:92241678; PMID:1572548
A:Accession: JC1145
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 18-109 <GRI>
A:Experimental source: salivary gland
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.6%; Score 7; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163
DB 4 ALLVLAL 10

RESULT 68
INR8
insulin precursor - rabbit
N:Alternate names: preproinsulin
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 23-Aug-1997 #text_change 18-Jun-1999
C:Accession: A53438; A01581
R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.R.; Menon, R.K.; Zahn, D.
A:Title: Insulin gene expression and insulin synthesis in mammalian neuronal cells.
A:Reference number: A53438; MUID:94179230; PMID:8132571
A:Accession: A53438
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-110 <DEV>
A:Cross-references: GB:U03610; NID:G467970; PIDN:AAA19033.1; PID:G467971
A:Smith, L.F.
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01581
A:Molecule type: protein
A:Residues: 25-54; 90-110 <SMI>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54,90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting C peptide #status predicted <CPDP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
|||
Db 10 LLALLVL 16

RESULT 69
IPRT1
insulin 1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 24-Sep-1999
C;Accession: A90788; A90789; A94331; B92120; I51945; A01589
R;Cordall, B.; Bell, G.; Tischer, E.; DenOto, F.M.; Ullrich, A.; Pictet, R.; Rutter, W.D.
Cell 18, 533-543, 1979
A;Title: Isolation and characterization of a cloned rat insulin gene.
A;Reference number: A90788; MUID:80045034; PMID:498283
A;Accession: A90788
A;Molecule type: DNA
A;Residues: 1-110 <COR>
A;Cross-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R;Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
Cell 18, 545-558, 1979
A;Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A;Reference number: A90789; MUID:80045035; PMID:498284
A;Accession: A90789
A;Molecule type: DNA
A;Residues: 1-110 <COM>
A;Cross-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R;Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliash, E.; Aten, B.; Oye
Recent Prog. Horm. Res. 25, 207-282, 1969
A;Title: Proinsulin and the biosynthesis of insulin.
A;Reference number: A94331; MUID:70067613; PMID:4311938
A;Accession: A94331
A;Molecule type: protein
A;Residues: 25-54; 90-110 <STE>
R;Tagger, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A;Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A;Reference number: A92120; MUID:73061498; PMID:4640931
A;Accession: B92120
A;Molecule type: protein
A;Residues: 57-87 <TAG>
R;Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-432, 1980
A;Title: The structure of rat preproinsulin genes.
A;Reference number: I51945; MUID:80240379; PMID:6249167
A;Accession: I51945
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-110 <RES>
A;Cross-references: GB:M25584; NID:G204947; PIDN:AAA41439.1; PID:G204948
C;Genetics:
A;Gene: INS1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54, 90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting peptide #status experimental <CEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96, 43-109, 95-100/Disulfide bonds: #status experimental

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
|||
Db 10 LLALLVL 16

RESULT 70

S30268
protein hdeA precursor - Escherichia coli (strain K-12)
N;Alternate names: 10K S protein precursor; hypothetical protein A
C;Species: Escherichia coli
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
C;Accession: S30268; S30266; S47730; A65149
R;Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A;Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A;Reference number: S30261; MUID:93204884; PMID:8455549
A;Accession: S30268
A;Molecule type: DNA
A;Residues: 1-110 <YOS>
A;Cross-references: EMBL:D11109; NID:G216429; PIDN:BA01883.1; PID:G216431
A;Accession: S30266
A;Molecule type: protein
A;Residues: 22-40 <YOW>
R;Plunkett, G.
submitted to the EMBL Data Library, March 1994
A;Reference number: S47666
A;Molecule type: DNA
A;Residues: 1-110 <PLU>
A;Cross-references: EMBL:U00039; NID:G466582; PIDN:AAB18486.1; PID:G466647
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A65149
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <BLAT>
A;Cross-references: GB:AE000427; GB:U00096; NID:G1789919; PIDN:AA076535.1; PID:G1789926
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: hdeA
C;Superfamily: Escherichia coli protein hdeA
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-110/Product: 10K-S protein #status experimental <MAT>

Query Match 0.6%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
|||
Db 9 LGGLLLL 15

RESULT 71
F91177
protein hdeA precursor - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
C;Accession: F91177
R;Hayaishi, T.; Wakino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91177
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA037813.1; PID:G13363864; GSPDB:GNO0154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECS4390
C;Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
 |||||
 Db 9 LGGLLLL 15

RESULT 72
 G86023
 protein hdeA precursor - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C:Accession: G86023
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 illet, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: G86023
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-110 <STO>
 A:Cross-references: GB:AB005174; NID:G12518219; PIDN:AAG58651.1; GSPDB:GN00145; UWGP:249
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: hdeA
 C:Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
 |||||
 Db 9 LGGLLLL 15

RESULT 73
 C95870
 hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymB
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: C95870
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar
 Proc. Natl. Acad. Sci. U.S.A. 98, 9899-9894, 2001
 A>Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21365608; PMID:11481431
 A:Accession: C95870
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-114 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48627.1; PID:G15140099; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 283, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaud, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C:Contents: annotation
 C:Genetics:
 A:Gene: Smb20236
 A:Genome: plasmid

Query Match 0.6%; Score 7; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 RLGLSLA 110

Db 88 RLGLSLA 94
 |||||

RESULT 74
 D75434
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75434
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.W.
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: D75434
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <WHI>
 A:Cross-references: GB:AB001962; GB:AE000513; NID:G6458855; PIDN:AAF10702.1; PID:G645886
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1125
 A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
 |||||
 Db 10 LLLALL 16

RESULT 75
 B81215
 conserved hypothetical inner membrane protein NMA2196 [imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
 C:Accession: B81215; G81792
 R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: B81215
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-132 <TET>
 A:Cross-references: GB:AE002386; GB:AE002098; NID:G7225512; PIDN:AAF40742.1; PID:G722551
 A:Experimental source: serogroup B, strain MC58
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jørgensen, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: G81792
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-132 <PAR>
 A:Cross-references: GB:AL162758; GB:AL157959; NID:G7380672; PIDN:CAB85407.1; PID:G738081
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMB0291; NMA2196

Query Match 0.8%; Score 7; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLA 1157

1ue sep 21 14:40:33 2004

Db 45 GGLLLA 51

Search completed: September 21, 2004, 13:42:32
Job time : 32 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:06:29 ; Search time 29 seconds
(without alignments)
2133.081 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188
Sequence: 1 MDLPRGLVVAWALSMPGFT.....FRSARRRREGLDTPFKVLE 1188

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911	76.7	1189	1	ITAH_HUMAN
2	15	1.3	285	1	ITAH_CHICK
3	15	1.3	1151	1	ITAH_HUMAN
4	15	1.3	1180	1	ITAH_RAT
5	13	1.1	1167	1	ITAG_HUMAN
6	11	0.9	1170	1	ITAG_BOVIN
7	11	0.9	1181	1	ITAG_HUMAN
8	9	0.8	312	1	CAH4_BOVIN
9	9	0.8	371	1	LEU3_SCHPO
10	9	0.8	1178	1	ITAG_MOUSE
11	8	0.7	76	1	CD24_MOUSE
12	8	0.7	76	1	CD24_RAT
13	8	0.7	80	1	CD24_HUMAN
14	8	0.7	97	1	LST1_HUMAN
15	8	0.7	103	1	FTSB_SALTY
16	8	0.7	105	1	FTSB_KLEAE
17	8	0.7	179	1	IL22_HUMAN
18	8	0.7	234	1	EMBP_CAVPO
19	8	0.7	264	1	COMT_RAT
20	8	0.7	402	1	OPDE_PSEAE
21	8	0.7	423	1	NSMA_HUMAN
22	8	0.7	444	1	ENGA_MYCPE
23	8	0.7	451	1	HISX_COREF
24	8	0.7	467	1	DADR_HUMAN
25	8	0.7	467	1	SIL7_HUMAN
26	8	0.7	470	1	CPBK_MOUSE
27	8	0.7	487	1	YLIC_PSEPU
28	8	0.7	491	1	CPB1_RAT
29	8	0.7	491	1	CPB2_RAT
30	8	0.7	500	1	CPBA_MOUSE
31	8	0.7	604	1	VE1_EPV2
32	8	0.7	605	1	VE1_EPV1
33	8	0.7	811	1	TMS6_HUMAN

34	8	0.7	1050	1	ITAS_XENLA
35	8	0.7	1146	1	ITAI_DROME
36	8	0.7	1152	1	ITAM_HUMAN
37	8	0.7	1157	1	DD37_HUMAN
38	8	0.7	1162	1	ITAD_HUMAN
39	8	0.7	1163	1	ITAX_HUMAN
40	8	0.7	1173	1	TSPI_XENLA
41	8	0.7	1179	1	ITAE_HUMAN
42	8	0.7	1460	1	PMPC_CHLMU
43	8	0.7	1507	1	SET2_CABEL
44	8	0.7	1664	1	RPA1_YEAST
45	8	0.7	2238	1	RPL1_BUNYM
46	7	0.6	95	1	LST1_MOUSE
47	7	0.6	109	1	FEL2_FELCA
48	7	0.6	110	1	HDEA_ECOLI
49	7	0.6	110	1	INS1_RAT
50	7	0.6	110	1	INS_RABIT
51	7	0.6	139	1	IGF_MXGL
52	7	0.6	139	1	LAMP_PETMA
53	7	0.6	141	1	HBAD_CHICK
54	7	0.6	141	1	HBAD_MELGA
55	7	0.6	142	1	MX_CHICK
56	7	0.6	160	1	BIK_HUMAN
57	7	0.6	161	1	NUOE_AQUAE
58	7	0.6	161	1	TCH2_ABATH
59	7	0.6	171	1	YF87_METUA
60	7	0.6	176	1	YD2F_SCHPO
61	7	0.6	194	1	HIS7_CHLITE
62	7	0.6	201	1	2239_MOUSE
63	7	0.6	206	1	GPBB_MOUSE
64	7	0.6	208	1	PSMB_SULSO
65	7	0.6	213	1	KETA_PYRAE
66	7	0.6	214	1	CH13_HUMAN
67	7	0.6	214	1	CYB_CERCE
68	7	0.6	220	1	FGF3_CHICK
69	7	0.6	221	1	SDFL_HUMAN
70	7	0.6	223	1	EMBP_MOUSE
71	7	0.6	227	1	EMBP_RAT
72	7	0.6	232	1	EMBI_CAVPO
73	7	0.6	235	1	YBEU_ECOLI
74	7	0.6	235	1	YOGA_ECOLI
75	7	0.6	236	1	RR2_MAIZE
76	7	0.6	236	1	RR2_ORYSA
77	7	0.6	236	1	RR2_WHEAT
78	7	0.6	240	1	IBP6_HUMAN
79	7	0.6	253	1	GX42_RAT
80	7	0.6	253	1	Y685_METUA
81	7	0.6	260	1	COX3_ASTPE
82	7	0.6	260	1	COX3_DINSE
83	7	0.6	261	1	COX3_PARLI
84	7	0.6	261	1	COX3_BRALA
85	7	0.6	262	1	COX3_METSE
86	7	0.6	263	1	LPSA_PASPA
87	7	0.6	270	1	REP8_HUMAN
88	7	0.6	276	1	NOSY_PSEST
89	7	0.6	276	1	UPK_BACSU
90	7	0.6	295	1	Y234_METUA
91	7	0.6	301	1	PANE_PYRHO
92	7	0.6	306	1	COAA_STRMU
93	7	0.6	309	1	MAT1_HUMAN
94	7	0.6	309	1	NTCB_SYNP7
95	7	0.6	311	1	MRAP_CAUCR
96	7	0.6	311	1	PYDA_ENTFA
97	7	0.6	312	1	O8H2_HUMAN
98	7	0.6	313	1	LSU3_NEILA
99	7	0.6	313	1	CAU1_HUMAN
100	7	0.6	315	1	CAU1_HUMAN

ALIGNMENTS

RESULT 1

ITAH_HUMAN
 ID ITAH_HUMAN STANDARD; PRT; 1189 AA.
 AC Q9UKX5; Q9UKQ1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-11 precursor.
 GN ITG11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal heart, and Osteoblast;
 RX MEDLINE=94117678; PubMed=10486209;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
 RA Wang S.-X., Morris C.M., Krissansen G.W.;
 RT "Cloning, sequence analysis, and chromosomal localization of the novel
 human integrin alpha11 subunit (ITG11).";
 RL Genomics 60:179-187(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal muscle, and Uterus;
 RX MEDLINE=9395147; PubMed=10464311;
 RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
 RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
 Integrin. A collagen-binding, i domain-containing, beta(1)-associated
 integrin alpha-chain present in muscle tissues.";
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [3]
 RP SEQUENCE OF 954-1188 FROM N.A.
 RC TISSUE=Fibroblast;
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
 CC ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
 CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
 CC -1- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
 CC FETAL MUSCLE CELLS (IN VITRO).
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF109681; AAF01258.1; -;
 CC EMBL; AF130278; AAD51919.2; -;
 CC EMBL; AL359064; CAB94392.1; -;
 CC HSSP; PI7301; IA0X.
 CC
 CC Genew; HGNC:6136; ITG11.
 CC MIM; 604789; -;
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC GO; GO:0007517; P:muscle development; TAS.
 CC InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRIN.
 DR PRINTS; P300453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; vwa; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 1189
 FT DOMAIN 23 1142
 FT TRANSMEM 1143 1165
 FT DOMAIN 1166 1189
 FT REPEAT 38 94
 FT REPEAT 102 163
 FT DOMAIN 167 345
 FT REPEAT 359 420
 FT REPEAT 422 475
 FT REPEAT 477 537
 FT REPEAT 539 598
 FT REPEAT 601 653
 FT DOMAIN 1154 1162
 FT DOMAIN 1174 1177
 FT CA_BIND 488 496
 FT CA_BIND 551 559
 FT CA_BIND 613 621
 FT DISULFID 76 83
 FT DISULFID 121 139
 FT DISULFID 129 159
 FT DISULFID 659 668
 FT DISULFID 674 729
 FT DISULFID 781 787
 FT DISULFID 881 893
 FT CARBOHYD 82 82
 FT CARBOHYD 95 95
 FT CARBOHYD 291 291
 FT CARBOHYD 331 331
 FT CARBOHYD 358 358
 FT CARBOHYD 449 449
 FT CARBOHYD 462 462
 FT CARBOHYD 528 528
 FT CARBOHYD 642 642
 FT CARBOHYD 694 694
 FT CARBOHYD 857 857
 FT CARBOHYD 894 894
 FT CARBOHYD 973 973
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1040 1040
 FT VARIANT 433 433
 FT VARIANT 524 524
 FT VARIANT 972 972
 FT VARIANT 1003 1003
 FT VARIANT 1030 1030
 FT VARIANT 1094 1094
 FT VARIANT 1094 1094
 FT SEQUENCE 1189 AA; 133609 MW; 60303C08A4A4CD52 CRC64;
 Query Match 76.7%; Score 911; DB 1; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVWAWALSHPGTDTFNMDTRKPRVPGSRFAFFGYTQQCHISGNKMLVGA 60
 Db 1 MDLPRGLVWAWALSHPGTDTFNMDTRKPRVPGSRFAFFGYTQQCHISGNKMLVGA 60

QY	61	PLETNGYQKTDGVYKCPVIHNGCTKLNLRVTLSNVSEKDNMRGLSLATNPKDNSFLA	120	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
DB	61	PLETNGYQKTDGVYKCPVIHNGCTKLNLRVTLSNVSEKDNMRGLSLATNPKDNSFLA	120	OC	Gallus.
QY	121	CSPLWSHECGSSYTTTGMCSRNSNPRFSKTVAPALQRCQTYNDIVIVLDGNSNIYPWVE	180	RN	NCBI_TaxID=9031;
DB	121	CSPLWSHECGSSYTTTGMCSRNSNPRFSKTVAPALQRCQTYNDIVIVLDGNSNIYPWVE	180	RP	[1]
QY	181	VOHFLNLIKKEFVIGPQIQGVVQYGEDVWVEHFLNDYRSVKDVVEAAASHIEQRGGTET	240	RC	SEQUENCE FROM N.A.
DB	181	VOHFLNLIKKEFVIGPQIQGVVQYGEDVWVEHFLNDYRSVKDVVEAAASHIEQRGGTET	240	TX	TISSUE=Gizzard;
QY	241	RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPLEKVIQOESRDNTRYAVAVL	300	RA	MEDLINE=94357930; PubMed=7521332;
DB	241	RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPLEKVIQOESRDNTRYAVAVL	300	RT	Kern A., Brissewitz R., Bank I., Marcantonio E.B.;
QY	301	GYNRRGINPETELNEIKYIASDPDKHFNVTDEALKDIDVALGDRIFSLEGTKNET	360	RL	"The role of the I domain in ligand binding of the human integrin
DB	301	GYNRRGINPETELNEIKYIASDPDKHFNVTDEALKDIDVALGDRIFSLEGTKNET	360	CC	alpha 1 beta 1.";
QY	361	SFGELEMSQTFSSHVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPEELKN	420	CC	J Biol. Chem. 269:22811-22816(1994).
DB	361	SFGELEMSQTFSSHVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPEELKN	420	CC	CC
QY	421	HGAYLGYTVTSVSSRQGRVYVAGAPRHNHTGKVILFTWHNRSLSLTIHOAMRQOIGSYF	480	CC	CC
DB	421	HGAYLGYTVTSVSSRQGRVYVAGAPRHNHTGKVILFTWHNRSLSLTIHOAMRQOIGSYF	480	CC	CC
QY	481	GSEITSDVDIGDGTVDLLVAGAPMYFNEGRERKVVYELRQRFYNGTLDKSHSYQNA	540	CC	CC
DB	481	GSEITSDVDIGDGTVDLLVAGAPMYFNEGRERKVVYELRQRFYNGTLDKSHSYQNA	540	CC	CC
QY	541	RFSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQRTASBLATG	600	CC	CC
DB	541	RFSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQRTASBLATG	600	CC	CC
QY	601	LOVFGCSIHQDLNEDGLDLAVGALGNVILWSRPVQVINALSHFEPSKINIFHRDCK	660	CC	CC
DB	601	LOVFGCSIHQDLNEDGLDLAVGALGNVILWSRPVQVINALSHFEPSKINIFHRDCK	660	CC	CC
QY	661	RSRGDACLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV	720	CC	CC
DB	661	RSRGDACLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV	720	CC	CC
QY	721	LLSSGQELCRINPHVLDADYKVPVTFSEYSLDHPGMLDDGPTTLRVSVPFWNG	780	CC	CC
DB	721	LLSSGQELCRINPHVLDADYKVPVTFSEYSLDHPGMLDDGPTTLRVSVPFWNG	780	CC	CC
QY	781	CNEDEHCVPLDILVARSDDLPTAMEYCORVLRKPAQDCSAYTLGFDITTVFIESTRQVAV	840	CC	CC
DB	781	CNEDEHCVPLDILVARSDDLPTAMEYCORVLRKPAQDCSAYTLGFDITTVFIESTRQVAV	840	CC	CC
QY	841	EATLENRGENAYSTVLNISOSANLOFASLIQKEDSDGSIQVNEERLQKQVNCVYPPFF	900	CC	CC
DB	841	EATLENRGENAYSTVLNISOSANLOFASLIQKEDSDGSIQVNEERLQKQVNCVYPPFF	900	CC	CC
QY	901	RAKAKVAFRLD 911		CC	CC
DB	901	RAKAKVAFRLD 911		CC	CC
RESULT 2					
ITAL_CHICK					
ID	ITAL_CHICK	STANDARD;	PRT;	285 AA.	
AC	Q90615;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a)				
DE	(Fragment).				
GN	ITGAL.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

QY	164	DIVIVLDGNSNIYPW	178	Query Match	1.3%; Score 15; DB 1; Length 285;
DB	63	DIVIVLDGNSNIYPW	77	Best Local Similarity	100.0%; Pred.No. 1.8e-06;
					Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 3					
ITAL_HUMAN					
ID	ITAL_HUMAN	STANDARD;	PRT;	1151 AA.	
AC	P56199;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a)				
GN	ITGAL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				

RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit.";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PRO; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1QC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGA1.
DR MIM; 192968;
DR GO; GO:0008305; C: integrin complex; TAS.
DR GO; GO:0004895; F: cell adhesion receptor activity; NAS.
DR GO; GO:0005118; F: collagen binding; TAS.
DR GO; GO:0007160; P: cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113
FT TRANSMEM 1114 1136
FT DOMAIN 1137 1151
FT REPEAT 16 75
FT REPEAT 2 75
FT DOMAIN 147 360
FT REPEAT 349 404
FT REPEAT 405 457
FT REPEAT 459 520
FT REPEAT 540 599
FT REPEAT 602 654
FT CA_BIND 470 478
FT CA_BIND 552 560
FT CA_BIND 614 622
FT SITE 1139 1142
FT DISULFID 54 64
FT DISULFID 660 669
FT DISULFID 675 728
FT DISULFID 780 786
FT DISULFID 850 858
FT DISULFID 1002 1034
FT DISULFID 1037 1044
FT CARBOHYD 46 46
FT CARBOHYD 72 72
FT CARBOHYD 77 77
FT CARBOHYD 84 84
FT CARBOHYD 189 189
FT CARBOHYD 289 289
FT CARBOHYD 313 313
FT CARBOHYD 374 374
FT CARBOHYD 390 390
FT CARBOHYD 432 432
FT CARBOHYD 504 504
FT CARBOHYD 671 671
FT CARBOHYD 720 720
FT CARBOHYD 752 752

FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3CIAAF52808 CRC64;
Query Match 1.3%; Score 15; DB 1; Length 1151;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVVLGSGNSIYRW 178
DB 144 DIVVLGSGNSIYRW 158
RESULT 4
ID ITAL RAT
AC P18614; STANDARD; PRT; 1180 AA.
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGA1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch P., Carbonetto S., Reichardt L.F.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen.";
RL J. Cell Biol. 111:709-720(1990).
RN [2]
RX X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliansky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alpha1beta1 integrin I-domain: insights into
integrin I-domain function.";
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC

EMBL; X52140; CAA36384.1; -.
 PIR; A35854.
 PDB; 1CK4; O3-MAY-00.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR002035; VWF_A.
 Pfam; PFO1839; FG-GAP; 3.
 Pfam; PFO0357; Integrin_A; 1.
 Pfam; PFO0092; vwa; 1.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWF; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; VWFA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Repeat; Calcium; Magnesium; 3D-structure.
 SIGNAL 1 28
 CHAIN 29 1180 INTEGRIN ALPHA-1.
 DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 1143 1165 POTENTIAL.
 DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
 REPEAT 44 103 FG-GAP 1.
 REPEAT 44 103 FG-GAP 2.
 REPEAT 175 388 VWFA.
 DOMAIN 175 388
 REPEAT 377 432 FG-GAP 3.
 REPEAT 433 484 FG-GAP 4.
 REPEAT 485 565 FG-GAP 5.
 REPEAT 567 626 FG-GAP 6.
 REPEAT 629 681 FG-GAP 7.
 REPEAT 681 707 POTENTIAL.
 CA_BIND 579 587 POTENTIAL.
 CA_BIND 641 649 GPFK MOTIF.
 SITE 1168 1172
 DISULFID 82 92 BY SIMILARITY.
 DISULFID 687 696 BY SIMILARITY.
 DISULFID 702 755 BY SIMILARITY.
 DISULFID 807 813 BY SIMILARITY.
 DISULFID 877 885 BY SIMILARITY.
 DISULFID 1029 1062 BY SIMILARITY.
 DISULFID 1066 1073 BY SIMILARITY.
 CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 779 779 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 1007 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 1084 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 1103 1103 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 1114 1114 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;
 Query Match 1.3%; Score 15; DB 1; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 5.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 DIVILVDGNSIYPW 178
 |||||
 Db 172 DIVILVDGNSIYPW 186
 |||||

FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1123 1145 POTENTIAL.
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT 38 97 FG-GAP 2.
 FT DOMAIN 167 350 VWFA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LEU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;

Query Match 1.1%; Score 13; DB 1; Length 1167;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 VIVLDGNSIYPW 178
 |||||
 Db 169 VIVLDGNSIYPW 181
 |||||

RESULT 6
 ID ITA2_BOVIN STANDARD; PRT; 1170 AA.
 AC P53710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
 GN ITGA2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=7511592;
 RA Kamata T., Puzon W., Takada Y.;
 RT Identification of putative ligand binding sites within I domain of
 RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663(1994).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE

CC CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC CC EXTRACELLULAR MATRIX.
 CC CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC CC ASSOCIATES WITH BETA-1.
 CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL; L25886; AAB59255.1; -.
 CC CC PIR; I45914; I45914.
 CC CC HSP; P17301; IAOX.
 CC CC InterPro; IPR000413; Integrin_alpha.
 CC CC InterPro; IPR002035; VWFA.
 CC CC Pfam; PF01839; FG-GAP; 3.
 CC CC Pfam; PF00357; Integrin_A; 1.
 CC CC Pfam; PF00092; vwa; 1.
 CC CC SMART; SM00191; Int_alpha; 5.
 CC CC SMART; SM00327; VWFA_1.
 CC CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC CC PROSITE; PS0234; VWFA; 1.
 CC CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC CC Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
 CC CC NON_TER 1 1
 CC CC SIGNAL <1 18 POTENTIAL.
 CC CC CHAIN 19 1170 INTEGRIN ALPHA-2.
 CC CC DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).
 CC CC TRANSMEM 1122 1143 POTENTIAL.
 CC CC DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).
 CC CC REPEAT 34 92 FG-GAP 1.
 CC CC REPEAT ? ? FG-GAP 2.
 CC CC DOMAIN 177 367 VWFA.
 CC CC REPEAT ? ? FG-GAP 3.
 CC CC REPEAT 423 475 FG-GAP 4.
 CC CC REPEAT 477 538 FG-GAP 5.
 CC CC REPEAT 540 599 FG-GAP 6.
 CC CC REPEAT 604 656 FG-GAP 7.
 CC CC CA_BIND 488 496 POTENTIAL.
 CC CC CA_BIND 552 560 POTENTIAL.
 CC CC CA_BIND 616 624 POTENTIAL.
 CC CC SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).
 CC CC SITE 1146 1150 GFFR MOTIF.
 CC CC DISULFID 72 81 BY SIMILARITY.
 CC CC DISULFID 669 726 BY SIMILARITY.
 CC CC DISULFID 778 784 BY SIMILARITY.
 CC CC DISULFID 854 865 BY SIMILARITY.
 CC CC DISULFID 1008 1039 BY SIMILARITY.
 CC CC DISULFID 1044 1049 BY SIMILARITY.
 CC CC CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC VARIANT 580 580 R -> V.
 CC CC VARIANT 588 588 R -> K.
 CC CC VARIANT 725 725 R -> S.
 CC CC SEQUENCE 1170 AA; 128929 MW; BECEFLC5F2448FE1 CRC64;

Query Match 0.9%; Score 11; DB 1; Length 1170;
 Best Local Similarity 100.0%; Pred No. 0.05;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLGAPMY 505
 |||||
 DB 495 TDVLLGAPMY 505

RESULT 7
 ITA2 HUMAN STANDARD; PRT; 1181 AA.
 ID ITA2 HUMAN
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN ITGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
 RP TISSUE=Endothelial cells;
 RC MEDLINE=89308879; PubMed=2545729;
 RA Takada Y., Hemler M.E.;
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
 (platelet GPIa): homology to other integrins and the presence of a
 possible collagen-binding domain.";
 RL J. Cell Biol. 109:397-407 (1989).
 [2]
 RN SEQUENCE FROM N.A.
 RP Rieder M.J., Arnel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [3]
 RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
 RP MEDLINE=98013223; PubMed=9353312;
 RA Emley J., King S.L., Bergelson J.M., Liddington R.C.;
 RT "Crystal structure of the I domain from integrin alpha2beta1.";
 RL J. Biol. Chem. 272:28512-28517 (1997).
 [4]
 RN VARIANT HPA-5 (BR).
 RP MEDLINE=94043762; PubMed=7901236;
 RA Santoso S., Kalb R., Walka M., Kiefel V., Mueller-Eckhardt C.,
 RA Newman P.J.;
 RT "The human platelet alloantigens Br(a) and Brb are associated with a
 single amino acid polymorphism on glycoprotein Ia (integrin subunit
 alpha 2).";
 RL J. Clin. Invest. 92:2427-2432 (1993).
 [5]
 RN VARIANT GLU-534.
 RP MEDLINE=20206009; PubMed=10744142;
 RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
 gene polymorphism on coronary artery disease and acute myocardial
 infarction.";
 RL Thromb. Haemost. 83:392-396 (2000).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
 CC associates with beta-1. Interacts with HPS5.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific
 CC alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)
 CC has Glu-534. HPA-5B is involved in neonatal alloimmune
 CC thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a
 CC role in coronary artery disease (CAD).
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VMPA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X17033; CRA34894.1; -;
 CC EMBL: AF525556; AAM34795.1; -;
 CC PIR: A33998; A33998.
 CC PDB: 1A0X; 25-NOV-98.
 CC PDB: 1DZ1; 02-AUG-01.
 CC Genew: HGNC:6137; ITGA2.
 CC MIM: 192974; -;
 CC CO: GO:0008305; C:integrin complex; TAS.
 CC GO: GO:0005886; C:plasma membrane; TAS.
 CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO: GO:0005181; F:collagen binding; TAS.
 CC GO: GO:0007596; P:blood coagulation; TAS.
 CC GO: GO:0007160; P:cell-matrix adhesion; TAS.
 CC GO: GO:0007397; P:histogenesis and organogenesis; TAS.
 CC InterPro: IPR000413; Integrin_alpha.
 CC InterPro: IPR002035; VWF_A.
 CC Pfam: PF01839; FG-GAP; 3.
 CC Pfam: PF00357; Integrin_A; 1.
 CC Pfam: PF00092; vwa; 1.
 CC SMART: SM00191; Int_alpha; 5.
 CC SMART: SM00327; VWA; 1.
 CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE: PS00234; VMPA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
 KW 3D-structure. 1 29
 FT SIGNAL 1 29 INTEGRIN ALPHA-2.
 FT CHAIN 30 1181 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 30 1132 POTENTIAL.
 FT TRANSMEM 1133 1134 POTENTIAL.
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1155 1161 INTERACTION WITH HPS5.
 FT REPEAT 45 103 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT REPEAT ? ? VMPA.
 FT DOMAIN 188 378 FG-GAP 3.
 FT REPEAT 378 433 FG-GAP 4.
 FT REPEAT 434 486 FG-GAP 5.
 FT REPEAT 488 549 FG-GAP 6.
 FT REPEAT 551 610 FG-GAP 7.
 FT REPEAT 615 667 POTENTIAL.
 FT CA_BIND 499 507 POTENTIAL.
 FT CA_BIND 563 571 POTENTIAL.
 FT CA_BIND 627 635 POTENTIAL.
 FT SITE 1157 1161 GFGR MOTIF.
 FT DISULFID 83 92 BY SIMILARITY.
 FT DISULFID 680 737 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT DISULFID 865 876 BY SIMILARITY.
 FT DISULFID 1019 1050 BY SIMILARITY.
 FT DISULFID 1055 1060 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckert G., Aert R., Robben J., Grymoprez B.,
RA Welljens I., Vansteenkiste E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Porashkin J.,
RA Spakovski G.V., Usery D., Barrell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-4-hydroxy-4-
CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
CC oxopentanoate. The product decarboxylates to 4-methyl-2-
CC oxopentanoate.
CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.
CC -!- PATHWAY: Leucine biosynthesis; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate
CC dehydrogenases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36910; AAA35316.1; -.
CC EMBL; AL031174; CAA20106.1; -.
CC EMBL; T43407; T43407.
CC HSP; P12010; 2XQY.
CC GeneDB_Spombe; SPBC1A4.02c; -.
CC InterPro; IPR001804; IsoDH.
CC InterPro; IPR004429; LeuB.
CC Pfam; PF00180; IsoDH; 1.
CC TIGRFAMs; TIGR00169; leuB; 1.
CC PROSITE; PS00470; IDH_IMDH; 1.
CC Oxidoreductase; Leucine biosynthesis; NAD.
KW Oxidoreductase; Leucine biosynthesis; NAD.
SQ SEQUENCE 371 AA; 39732 MW; 65AAAE6AA94D45EE CRC64;

Query Match 0.8%; Score 9; DB 1; Length 371;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 DGVLGAVG 386
DB 69 DGVLGAVG 77

RESULT 10
ITR2 MOUSE
ID ITR2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1) SEQUENCE FROM N.A.
RN STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RC MEDLINE=94363406; PubMed=8081889;
RX Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RT Cell Adhes. Commun. 2:131-143(1994).
(2) SEQUENCE OF 450-1178 FROM N.A.
RN TISSUE=Lung;
RC MEDLINE=94355691; PubMed=7521231;
RX Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGEN, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS. ALPHA-2
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF A DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWF A domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z29987; CAA82877.1; -.
CC EMBL; X75427; CAA53178.1; -.
CC FIR; S44142; S44142.
CC HSP; P17301; LAOX.
CC MGD; MGI:96600; Itga2.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWF A; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178 INTEGRIN ALPHA-2.
FT DOMAIN 27 1129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1130 1151 POTENTIAL.
FT DOMAIN 1152 1178 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 100 FG-GAP 1.
FT REPEAT ? 2 FG-GAP 2.
FT DOMAIN 185 375 VWF A.
FT REPEAT ? 483 FG-GAP 3.
FT REPEAT 431 483 FG-GAP 4.
FT REPEAT 485 546 FG-GAP 5.
FT REPEAT 548 607 FG-GAP 6.
FT REPEAT 612 664 FG-GAP 7.
FT CA_BIND 496 504 POTENTIAL.
FT

```

FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 OFFER MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 1178;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAP 503
DB 503 TDVLLVGAP 511

RESULT 11
CD24 MOUSE
ID CD24 MOUSE STANDARD; PRT; 76 AA.
AC P24807; P26691;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (M1/69-J11D heat stable antigen)
DE (HSA) (Nectadrin) (115-52) (X62 heat stable antigen) (R13-AG).
GN CD24 OR CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90361906; PubMed=2118158;
RA Kay R., Takei F., Humphries R.K.;
RT "Expression cloning of a cDNA encoding M1/69-J11D heat-stable
RT antigens";
RL J. Immunol. 145:1952-1959 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA, and Swiss albino X BALB/c; TISSUE=Spleen;
RC MEDLINE=91209380; PubMed=2019286;
RA Wanger R.H., Avane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called
RT the heat-stable antigen";
RL Eur. J. Immunol. 21:1039-1046 (1991).
RN [3]
RP REVISIONS.
RA Nielsen P.J.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA; TISSUE=Spleen;
RC MEDLINE=94043127; PubMed=8226859;
RA Wanger R.H., Rochelle J.M., Seidin M.F., Koehler G., Nielsen P.J.;
RT "The heat stable antigen (mouse CD24) gene is differentially
RT regulated but has a housekeeping promoter.";
RL J. Biol. Chem. 268:23345-23352 (1993).
RN [5]
RP SEQUENCE OF 27-53.

```

```

RC STRAIN=C57BL/6;
RX MEDLINE=92412120; PubMed=1530634;
RA Hitsumoto Y., Nakano A., Ohnishi H., Hamada F., Saheki S.,
RA Takeuchi N.;
RT "Purification of the murine heat-stable antigen from erythrocytes.";
RL Biochem. Biophys. Res. Commun. 187:773-777 (1992).
CC -!- FUNCTION: May have a specific role to play in early thymocyte
CC development.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: In lymphoid, myeloid, and erythroid cells.
CC -!- PTM: Extensively O-glycosylated (By similarity).
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M58661; AAA39481.1; -
DR EMBL; X56469; CAA39841.1; -
DR EMBL; X72910; CAA51415.1; -
DR EMBL; X53825; CAA37822.1; -
DR PIR; A43537; A43537
DR MGI; MGI:88323; Cd24a.
KW Glycoprotein; GPI-anchor; Membrane; Signal; Antigen; Lipoprotein.
FT SIGNAL 1 26
FT CHAIN 27 53
FT PROPEP 54 76 SIGNAL TRANSDUCER CD24.
FT CARBOHYD 27 27 REMOVED IN MATURE FORM (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .).
FT CARBOHYD 39 39 O-LINKED (PROBABLE).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .).
FT LIPID 53 53 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 76 AA; 7797 MW; 6853F121B33625EB CRC64;

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 12 GLLLLALL 19

RESULT 12
CD24 RAT
ID CD24 RAT STANDARD; PRT; 76 AA.
AC Q07490;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (Heat stable antigen) (HSA)
DE (Nectadrin).
GN CD24A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Embryonic brain;
RC MEDLINE=94122434; PubMed=8292828;
RA Shirasawa T., Akashi T., Sakamoto K., Takahashi H., Maruyama N.,
RA Hirokawa K.;
RT "Gene expression of CD24 core peptide molecule in developing brain
RT and developing non-neural tissues.";
RL Dev. Dyn. 198:1-13 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Fischer;

```

Key R., Rosten P.M., Humphries R.K.;
"CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor";
J. Immunol. 147:1412-1416(1991).
(2)
SEQUENCE FROM N.A., AND VARIANT VAL-57.
MEDLINE=93007871; PubMed=1327504;
Jackson D., Wabel R., Weber E., Bell J., Stahel R.A.;
"CD24, a signal-transducing molecule expressed on human B cells, is a major surface antigen on small cell lung carcinomas";
Cancer Res. 52:5264-5270(1992).
(3)
SEQUENCE FROM N.A.
TISSUE=Ovary;
MEDLINE=4238257; PubMed=12477932;
Straussberg R.L., Reingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Garavito P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
(4)
SEQUENCE OF 1-76 FROM N.A.
MEDLINE=95048364; PubMed=7959762;
Hough M.R., Rosten P.M., Sexton T.L., Kay R., Humphries R.K.;
"Mapping of CD24 and homologous sequences to multiple chromosomal loci";
Genomics 22:154-161(1994).
CC -!- FUNCTION: Modulates B-cell activation responses. Signaling could be triggered by the binding of a lectin-like ligand to the CD24 carbohydrates, and transduced by the release of second messengers derived from the GPI-anchor. Promotes AG-dependent proliferation of B-cells, and prevents their terminal differentiation into antibody-forming cells.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: B-cells.
CC -!- INDUCTION: Expression is lost when primary B-cells are induced to differentiate in antibody-forming cells.
CC -!- PTM: Extensively O-glycosylated.
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD24 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd4.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M58664; AAA35665.1; --
DR EMBL; X69397; CAA49195.1; --
DR EMBL; L33930; AAB58807.1; --
DR EMBL; BC07674; AA030767.1; --
DR EMBL; S75311; AAD14170.1; ALT_INIT.
DR PIR; I56114; A48996.
DR Genew; HGNC:1645; CD24.

DR MIM: 600074; -- P:humoral immune response; TAS.
 DR GO:0006959; GPI-anchor; Membrane; Signal; Polymorphism; Lipoprotein.
 KW Glycoprotein; GPI-anchor; Membrane; Signal; Polymorphism; Lipoprotein.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 59 SIGNAL TRANSDUCER CD24.
 FT PROPEP 60 80 REMOVED IN MATURE FORM (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 59 59 GPI-anchor amidated glycine (Potential).
 FT VARIANT 57 57 A -> V (in dbSNP:8734).
 FT /FTid=VAR.016156.
 FT CONFLICT 12 12 G -> W (IN REF. 4).
 FT CONFLICT 44 44 S -> T (IN REF. 4).
 FT SEQUENCE 80 AA; 8083 MW; DB1988B6808F833F CRC64;
 SQ
 Query Match 0.7%; Score 8; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1152 GLLLLALL 1159
 DB 12 GLLLLALL 19
 RESULT 14
 LST1_HUMAN STANDARD; PRT: 97 AA.
 ID LST1_HUMAN STANDARD; PRT: 97 AA.
 AC 000453; 000452; 000454; Q13669; Q9UUR5; Q9UUR6; Q9UUR7; Q9UUR8;
 AC Q9UUR5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Leukocyte specific transcript 1 protein (B144 protein).
 GN LST1 OR B144
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORM 3), TISSUE SPECIFICITY, AND INDUCTION.
 RC TISSUE=Blood;
 RX MEDLINE=86006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kiek G., Zwierzina H.,
 RA Weiss E.H.;
 RA "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4), TISSUE SPECIFICITY, AND
 INDUCTION.
 RP TISSUE=Lymphoblast;
 RC MEDLINE=98035883; PubMed=9367684;
 RX de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RA "Complex expression pattern of the TNF region gene LST1 through
 differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:594-600(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 5; 6; 7; 8 AND 9), FUNCTION, AND
 SUBCELLULAR LOCATION.
 RP TISSUE=Peripheral blood;
 RX MEDLINE=20171517; PubMed=10706707;
 RA Rollinger-Holzinger I., Sibl B., Pauly M., Griesser U., Hentges P.,
 RA Auer B., Pall G., Schratzberger P., Niederwieser D., Weiss E.H.,
 RA Zwierzina H.;
 RA "LST1: a gene with extensive alternative splicing and immunomodulatory
 function.";
 RL J. Immunol. 164:3169-3176(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;

RT "Sequence of the human major histocompatibility complex class III
 region.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=21372017; PubMed=11478849;
 RA Ragunathan A., Sivakamasundari R., Wolenski J., Poddar R.,
 RA Weissman M.;
 RL "Functional analysis of B144/LST1: a gene in the tumor necrosis
 factor cluster that induces formation of long filopodia in eukaryotic
 cells.";
 RT Exp. Cell Res. 268:230-244(2001).
 RL CC
 CC -!- FUNCTION: Possible role in modulating immune responses. Isoforms 1
 and 2 have an inhibitory effect on lymphocyte proliferation.
 CC Induces morphological changes including production of filopodia
 and microspikes when overexpressed in a variety of cell types and
 CC may be involved in dendritic cell maturation.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in
 a perinuclear region corresponding to the localization of the
 Golgi apparatus and throughout the cytoplasm.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=9;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1; Synonyms=LST1/A;
 CC IsoId=000453-1; Sequence=Displayed;
 CC Name=2; Synonyms=LST1/C;
 CC IsoId=000453-2; Sequence=VSP_050578;
 CC Name=3; Synonyms=plst1;
 CC IsoId=000453-3; Sequence=VSP_050579;
 CC Name=4; Synonyms=LST1/E;
 CC IsoId=000453-4; Sequence=VSP_050577;
 CC Name=5; Synonyms=LST1/K;
 CC IsoId=000453-5; Sequence=VSP_050584;
 CC Name=6; Synonyms=LST1/L;
 CC IsoId=000453-6; Sequence=VSP_050583, VSP_050587;
 CC Name=7; Synonyms=LST1/J;
 CC IsoId=000453-7; Sequence=VSP_050583, VSP_050586;
 CC Name=8; Synonyms=LST1/M;
 CC IsoId=000453-8; Sequence=VSP_050580;
 CC Name=9; Synonyms=LST1/N;
 CC IsoId=000453-9; Sequence=VSP_050581, VSP_050582;
 CC -!- TISSUE SPECIFICITY: Expressed in lung, tonsil, thymus, placenta,
 kidney, fetal spleen, fetal liver and brain.
 CC -!- INDUCTION: By interferon gamma.
 CC -!- SIMILARITY: Belongs to the LST1 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 EMBL: U00921; AAB57724.1; --
 DR EMBL: AF000424; AAB86998.1; --
 DR EMBL: AF000425; AAB86999.1; --
 DR EMBL: AF000426; AAB87000.1; --
 DR EMBL: Y18486; CAB59904.1; --
 DR EMBL: Y18487; CAB59905.1; --
 DR EMBL: Y18488; CAB60038.1; --
 DR EMBL: Y18489; CAB59906.1; --
 DR EMBL: Y18490; CAB59903.1; --
 DR EMBL: AF129756; AAD18090.1; --
 DR EMBL: AP000505; BAB63394.1; --
 DR Genew; HGNC:14189; LST1.
 DR MIN; 109170; --
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0009653; P:morphogenesis; NAS.

OS Klebsiella aerogenes.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=28451;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W70;
 RX MEDLINE=20566700; PubMed=11114933;
 RA Koiko M.M., Kapetanovich L.A., Lawrence J.G.;
 RT "Alternative pathways for siroheme synthesis in Klebsiella
 aerogenes.";
 RL J. Bacteriol. 183:328-335(2001).
 CC -!- FUNCTION: Required for the cell division process (By similarity).
 CC -!- SUBUNIT: May interact with ftsL (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
 (Potential). Colocalizes with ftsL to the division site (By
 similarity).
 CC -!- SIMILARITY: Belongs to the ftsB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF308466; AAG42461.1; -
 DR HANAP; MF_00599; -; 1
 DR InterPro; IPR007060; DivIC.
 DR Pfam; PF04977; DivIC; 1.
 KW Cell division; Transmembrane; Inner membrane; Coiled coil.
 FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 4 21 POTENTIAL.
 FT DOMAIN 22 105 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 31 74 COILED COIL (POTENTIAL).
 SQ SEQUENCE 105 AA; 11951 MW; 530471363FD3112A CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALLLV 1160
 Db 6 LLLALLLV 13
 RESULT 17
 IL22 HUMAN
 ID IL22 HUMAN STANDARD; PRT; 179 AA.
 AC Q9GZK6; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Interleukin-22 precursor (IL-22) (IL-10-related T-cell-derived
 DE inducible factor) (IL-TIF) (UNQ3099/PRO10096).
 GN IL22 OR ILTIF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20420346; PubMed=10954742;
 RA Dumoutier L., Van Roost B., Coliau D., Renaud J.-C.;
 RT "Human interleukin-10-related T cell-derived inducible factor:
 RT molecular cloning and functional characterization as an hepatocyte-
 RT stimulating factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10144-10149(2000).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21069354; PubMed=11197690;
 RA Dumoutier L., Van Roost B., Coliau D., Ameye G., Michaux L.,

RA Renaud J.-C.;
 RT "IL-TIF/IL-22: genomic organization and mapping of the human and mouse
 RT genes.";
 RL Genes Immun. 1:488-494(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20469498; PubMed=10875937;
 RA Xie M.-H., Aggarwal S., Ho W.-H., Foster J., Zhang Z., Stinson J.,
 RA Wood W.I., Goddard A.D., Gurney A.L.;
 RT "Interleukin (IL)-22, a novel human cytokine that signals through the
 RT interferon receptor-related proteins CRP2-4 and IL-22R.";
 RL J. Biol. Chem. 275:31335-31339(2000).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-158.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22887296; PubMed=12975309;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seehagiri S., Simmons L., Smith J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiscand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 CC -!- FUNCTION: Cytokine that contributes to the inflammatory response
 CC in vivo.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the IL-10 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ277247; CAC06085.1; -
 DR EMBL; AJ277248; CAC19409.1; -
 DR EMBL; AF279437; AAG22064.1; -
 DR EMBL; AF387519; AAK62468.1; -
 DR EMBL; AY358890; AAG89249.1; -
 DR Genew; HGNC:14900; IL22.
 DR MIM; 605330; -
 DR GO; GO:0005576; C:extracellular; IC.
 DR GO; GO:0045518; F:interleukin-22 receptor binding; NAS.
 DR GO; GO:0006953; P:acute-phase response; NAS.
 DR GO; GO:0007267; P:cell-cell signaling; IC.
 DR GO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR000098; Interleukin 10.
 DR PROSITE; PS00520; INTERLEUKIN 10; 1.
 KW Cytokine; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 179 INTERLEUKIN-22.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 158 158 S -> G.
 FT
 SQ SEQUENCE 179 AA; 20011 MW; 3C35E64D60CF9767 CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 179;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1153 LLLLLLV 1160
Db 21 LLLLLLV 28

RESULT 18
EMB2_CAVPO STANDARD; PRT; 234 AA.
ID ENB2_CAVPO
AC P35709;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein 2 precursor (MBP-2).
GN MBP2.
OS Cavia porcellus (Guinea pig).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hysticognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;

[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Eosinophil;
RX MEDLINE=91224343; PubMed=2026266;
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA Saito T., Tasaka K.;
RT "Comparison of the amino acid and nucleotide sequences between human
RT and two guinea pig major basic proteins.";
RL FEBS Lett. 282:56-60(1991).
CC -!- FUNCTION: MBP may play some important roles in the allergic
CC reactions and inflammations, since MBP is capable of releasing
CC histamine from mast cells and damaging the epithelial cells of
CC bronchial tubes. Antiparasitic and antibiotic.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00817; BAA00697.1; -.
CC PIR; S15102; S15102.
CC HSRP; P13727; I18U.
CC InterPro; IPR002352; Emaior basic.
CC InterPro; IPR001304; Lectin_C.
CC Pfam; PF00059; lectin; 1.
CC PRINTS; PR00770; EMAJORBASICP.
CC SMART; SM00034; CLECT; 1.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS00041; C-TYPE LECTIN_2; 1.
CC Eosinophil; Signal; Immune response; Antibiotic; Lectin;
CC Multigene family; Glycoprotein.
CC SIGNAL 1 15 POTENTIAL.
CC FT PROPEP 16 115 ACIDIC.
CC FT CHAIN 116 234 BOSINOPHIL GRANULE MAJOR BASIC PROTEIN 2.
CC FT DOMAIN 133 234 C-TYPE LECTIN (SHORT FORM).
CC FT DISULFID 135 232 BY SIMILARITY.
CC FT DISULFID 209 224 BY SIMILARITY.
CC FT CARBOHYD 69 69 O-LINKED (GLYCOSAMINGLYCAN) (BY
CC SIMILARITY).
CC
CC SEQUENCE 234 AA; 26140 MW; 7D926A942BF5116F CRC64;
Query Match 0.7%; Score 8; DB 1; Length 234;
Best Local Similarity 100.0%; Pred.No.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
Db 4 LLLLLLV 11

RESULT 19
COMT_RAT STANDARD; PRT; 264 AA.
ID COMT_RAT
AC P22734;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Catechol O-methyltransferase, membrane-bound form (EC 2.1.1.6)
DE (MB-COMT) [Contains: Catechol O-methyltransferase, soluble form
DE (S-COMT)].
DN COMT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94107221; PubMed=8280056;
RA Tenhunen J., Ullmanen I.;
RT "Production of rat soluble and membrane-bound catechol O-
RT methyltransferase forms from bifunctional mRNAs.";
RL Biochem. J. 296:595-600(1993).
RN [2]
RP SEQUENCE OF 11-264 FROM N.A.
RX MEDLINE=91033034; PubMed=2227437;
RA Salminen M., Lundstroem K., Tilgmann C., Savolainen R., Kalkkainen N.,
RA Ullmanen I.;
RT "Molecular cloning and characterization of rat liver catechol-O-
RT methyltransferase.";
RL Gene 93:241-247(1990).
RN [3]
RP SEQUENCE OF 1-10 FROM N.A., AND CHARACTERIZATION OF THE TWO FORMS.
RX MEDLINE=92111472; PubMed=1765063;
RA Ullmanen I., Lundstroem K.;
RT "Cell-free synthesis of rat and human catechol O-methyltransferase.
RT Insertion of the membrane-bound form into microsomal membranes in
RT vitro.";
RL Eur. J. Biochem. 202:1013-1020(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF SOLUBLE FORM.
RX MEDLINE=94173341; PubMed=8127373;
RA Vidgren J., Svensson L.A., Liljas A.;
RT "Crystal structure of catechol O-methyltransferase.";
RL Nature 368:354-358(1994).
CC -!- FUNCTION: Catalyzes the O-methylation, and thereby the
CC inactivation, of catecholamine neurotransmitters and catechol
CC hormones. Also shortens the biological half-lives of certain
CC neuroactive drugs, like L-DOPA, alpha-methyl DOPA and
CC isoproterenol.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + catechol = S-
CC adenosyl-L-homocysteine + guaiacol.
CC -!- COFACTOR: Magnesium.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (isoform S-COMT). Type II
CC -!- membrane protein (isoform MB-COMT).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, Membrane-bound/MB-COMT (shown here) and
CC Soluble/S-COMT, are produced by alternative initiation;
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO OTHER MAMMALIAN CATECHOL-O-METHYLTRANSFERASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z12651; CAA78276.1; -.

```

DR EMBL; M60754; AAA40882.1; ALT INIT.
 DR EMBL; M60753; AAA40881.1; ALT_INIT.
 DR PIR; S22090; S22090.
 DR PDB; 1VID; 11-JUL-96.
 DR InterPro; IPR002935; Methyltransf_3.
 DR InterPro; IPR000851; SAM_bind.
 DR Pfam; PF01596; Methyltransf_3; 1.
 KW Transferase; Methyltransferase; Neurotransmitter degradation;
 KW Catecholamine metabolism; Transmembrane; Signal-anchor; Magnesium;
 KW Alternative initiation; 3D-structure.
 FT CHAIN 1 264
 FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
 FT MEMBRANE-BOUND.
 FT CHAIN 45 264
 FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
 FT SOLUBLE.
 FT INIT MET 44 44
 FT TRANS MEM 3 19
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT ACT SITE 187 187
 FT ACT_SITE 242 242
 FT METAL 184 184
 FT METAL 212 212
 FT METAL 213 213
 FT METAL 48 59
 FT HELIX 62 63
 FT TURN 65 78
 FT HELIX 79 79
 FT TURN 86 100
 FT STRAND 104 108
 FT TURN 111 112
 FT HELIX 114 120
 FT TURN 121 122
 FT TURN 125 126
 FT STRAND 128 133
 FT HELIX 136 149
 FT TURN 150 150
 FT TURN 152 154
 FT STRAND 155 159
 FT HELIX 162 165
 FT HELIX 166 168
 FT HELIX 169 173
 FT STRAND 178 183
 FT HELIX 187 189
 FT TURN 190 199
 FT TURN 200 201
 FT STRAND 203 211
 FT HELIX 214 218
 FT HELIX 220 228
 FT TURN 230 231
 FT STRAND 232 240
 FT TURN 242 243
 FT STRAND 247 255
 SQ SEQUENCE 264 AA; 29597 MW; F535DF49C062854 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 264;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
 |||||
 Db 9 GLLLLALL 16

RESULT 20

OPDE_PSEAE STANDARD; PRT; 402 AA.
 AC C01602;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Transcription regulatory protein opde.

GN OPDE OR PA2219.

OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1 / H103;
 RX MEDLINE=93051258; PubMed=1427017;
 RA Huang H., Siehnel R.J., Bellido F., Rawling E., Hancock R.E.W.;
 RT "Analysis of two gene regions involved in the expression of the
 imipenem-specific, outer membrane porin protein OpdP of Pseudomonas
 aeruginosa";
 RT FEMS Microbiol. Lett. 76:267-274 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Hickey M.J., Brinkman F.S.D., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 opportunistic pathogen";
 RT Nature 406:959-964 (2000).
 CC -!- FUNCTION: REGULATES THE EXPRESSION OF OPD WHICH ENCODES THE
 IMIPENEM-SPECIFIC PORIN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO B.SUBTILIS IPA-79D.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z14064; CAA78446.1; -.
 DR EMBL; AE004648; AAG05607.1; -.
 DR PIR; S23860; S23860.
 DR InterPro; IPR007114; MPS.
 DR InterPro; IPR005828; Sub_transporter.
 DR Pfam; PF00083; sugar_tr_1.
 DR PROSITE; PSS0850; MFS; 1.
 KW Transcription regulation; Transmembrane; Complete proteome.
 FT TRANS MEM 22 42
 FT TRANS MEM 60 80
 FT TRANS MEM 86 106
 FT TRANS MEM 108 128
 FT TRANS MEM 147 167
 FT TRANS MEM 170 190
 FT TRANS MEM 220 240
 FT TRANS MEM 256 276
 FT TRANS MEM 296 316
 FT TRANS MEM 318 338
 FT TRANS MEM 348 368
 FT TRANS MEM 375 395
 FT TRANS MEM 402 AA; 41592 MW; OC5701C4AD2FDE16 CRC64;
 SQ SEQUENCE 402 AA; 41592 MW; OC5701C4AD2FDE16 CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 GSTLGLL 1154
 |||||
 Db 360 GSTLGLL 367

RESULT 21

NSMA HUMAN STANDARD; PRT; 423 AA.

AC O60956; Q9BWR3;

DT 28-FEB-2003 (Rel. 41, Created)

DR 28-FEB-2003 (Rel. 41, Last sequence update)
DR 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sphingomyelin phosphodiesterase 2 (EC 3.1.4.12) (Neutral
DE sphingomyelinase) (nSMase) (Lyso-platelet activating factor-
DE phospholipase C) (Lyso-PAF-PLC)
GN SMPD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RP MEDLINE=98188255; PubMed=9520418;
RX Tomiuk S., Hofmann K., Nix M., Zumbansen M., Stoffel W.;
RA "Cloned mammalian neutral sphingomyelinase: functions in sphingolipid
RT signaling?";
RL Proc. Natl. Acad. Sci. U.S.A. 95:3638-3643 (1998).
RN [2]
RN SEQUENCE FROM N.A.
RX TISSUE=Brain;
RC MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RN CHARACTERIZATION.
RX MEDLINE=20076490; PubMed=10608884;
RA Sawai H., Domae N., Nagan N., Hannun Y.A.;
RT "Function of the cloned putative neutral sphingomyelinase as
RT Lyso-platelet activating factor-phospholipase C";
RL J. Biol. Chem. 274:38133-38139 (1999).
CC -!- FUNCTION: Converts sphingomyelin to ceramide. Hydrolyze 1-acyl-2-
CC lyso-sn-glycero-3-phosphocholine (lyso-PC) and 1-O-alkyl-2-lyso-
CC sn-glycero-3-phosphocholine (lyso-platelet activating factor). The
CC physiological substrate seems to be Lyso-PAF.
CC -!- CATALYTIC ACTIVITY: Sphingomyelin + H₂O = N-acylsphingosine +
CC choline phosphate.
CC -!- COFACTOR: Magnesium.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- MISCELLANEOUS: This protein has an optimum pH of 6.5-7.5.
CC -!- SIMILARITY: Belongs to the neutral sphingomyelinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ222801; CAA10995.1; --
DR EMBL; BC000038; AAH00038.1; --
DR Genbank; HGNC:11121; SMPD2.
DR MIM; 603498; --
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0004767; F: sphingomyelin phosphodiesterase activity; TAS.

DR GO; GO:0006684; P: sphingomyelin metabolism; TAS.
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
KW Hydrolase; Transmembrane; Magnesium.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 354 374 POTENTIAL.
FT METAL 49 49 MAGNESIUM (BY SIMILARITY).
FT SITE 180 180 IMPORTANT FOR SUBSTRATE RECOGNITION (BY
FT SIMILARITY).
FT ACT_SITE 272 272 GENERAL BASE (BY SIMILARITY).
FT CONFLICT 3 3 L -> P (IN REF. 2).
SQ SEQUENCE 423 AA; 47592 MW; 05252A923E363171 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1152 GLLLLALL 1159
DB 334 GLLLLALL 341
RESULT 22
ENGA_MYCPE STANDARD; PRT; 444 AA.
ID ENGA_MYCPE
AC Q8EWH6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE GTP-binding protein engA.
DE ENGA OR MYPE2290.
GN Mycoplasma penetrans.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Iehikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300 (2002).
CC -!- FUNCTION: GTPase of unknown physiological role.
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. EngA subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF004170; BAC44020.1; --
DR HAMAP; MF_00195; -- 1.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR TIGRfams; TIGR00650; MG442; 2.
DR TIGRfams; TIGR00231; small_GTP; 2.
DR GTP-binding; Repeat; Complete proteome.
KW NP_BIND 8 15 GTP 1 (POTENTIAL).
FT NP_BIND 55 59 GTP 1 (POTENTIAL).
FT NP_BIND 118 121 GTP 1 (POTENTIAL).
FT NP_BIND 179 186 GTP 2 (POTENTIAL).
FT NP_BIND 226 230 GTP 2 (POTENTIAL).
FT NP_BIND 291 294 GTP 2 (POTENTIAL).
SQ SEQUENCE 444 AA; 51231 MW; A7B636F43D7CB1BE CRC64;

```

Query Match          0.7%; Score 8; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 IIVLDGS 172
DB 258 IIVLDGS 265

RESULT 23
HISX_COREF
ID HISX_COREF STANDARD; PRT; 451 AA.
AC Q8FNZ0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hispidinol dehydrogenase (SC 1.1.1.23) (HDH).
GN HISD OR CE2003.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ireo K.,
RA Gojochori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
CC -!- FUNCTION: Catalyzes the sequential NAD-dependent oxidations of L-
CC histidinol to L-histidinaldehyde and then to L-histidine (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: L-histidinol + 2 NAD(+) + H(2)O = L-histidine
CC + 2 NADH.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- PATHWAY: Histidine biosynthesis; ninth (last) step.
CC -!- SIMILARITY: Belongs to the hispidinol dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AP005220; BAC18813.1; -
CC HAMAP; MF 01024; -; 1.
CC InterPro; IPR001692; Histidinol dh.
CC Pfam; PF00815; Histidinol dh; 1.
CC PRINTS; PR00083; HOLDHGRNASE.
CC ProDom; PD002680; Histidinol dh; 1.
CC TIGRfams; TIGR00069; hisd; 1.
CC PROSITE; PS00611; HISOL_DEHYDROGENASE; FALSE NEG.
CC KW Histidine biosynthesis; Oxidoreductase; NAD; Metal-binding; Zinc;
CC Complete proteome.
CC FT ACT_SITE 332 BY SIMILARITY.
CC FT ACT_SITE 333 BY SIMILARITY.
CC FT METAL 263 ZINC (BY SIMILARITY).
CC FT METAL 266 ZINC (BY SIMILARITY).
CC FT METAL 366 ZINC (BY SIMILARITY).
CC FT METAL 425 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 451 AA; 47770 MW; 5A2D5AA7B90ECCF CRC64;

Query Match          0.7%; Score 8; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 334 DEAAALKDI 341
DB 402 DEAAALKDI 409

RESULT 24
D4DR_HUMAN
ID D4DR_HUMAN STANDARD; PRT; 467 AA.
AC P21917;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D(4) dopamine receptor (D(2C) dopamine receptor).
DR D4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ALLELE D4.7).
RX MEDLINE=92310588; PubMed=1319557;
RA van Tol H.H., Wu C.M., Guan H.-C., Ohara K., Bunzow J.R.,
RA Civelli O., Kennedy J., Seeman P., Niznik H.B., Jovanovic V.;
RT "Multiple dopamine D4 receptor variants in the human population."
RL Nature 358:149-152(1992).
RN [2]
RP SEQUENCE FROM N.A. (ALLELE D4.2).
RC TISSUE=Brain;
RX MEDLINE=91204054; PubMed=1840645;
RA van Tol H.H.M., Bunzow J.R., Guan H.-C., Sunahara R.K., Seeman P.,
RA Niznik H.B., Civelli O.;
RT "Cloning of the gene for a human dopamine D4 receptor with high
RT affinity for the antipsychotic clozapine."
RL Nature 350:610-614(1991).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=93038566; PubMed=1359063;
RA Livingston C.D., Strange P.G., Naylor L.H.;
RT "Molecular modelling of D2-like dopamine receptors."
RL Biochem. J. 287:277-282(1992).
RN [4]
RP VARIANT GLY-194.
RX MEDLINE=95243275; PubMed=7726213;
RA Seeman P., Ulpian C., Chouinard G., van Tol H.H.M., Dwosh H.,
RA Lieberman J.A., Sininovich K., Liu I.S.C., Wayne J., Voruganti P.,
RA Hudson C., Serjeant G.R., Masibay A.S., Seeman M.V.;
RT "Dopamine D4 receptor variant, D4-glycine-194, in Africans, but not in
RT Caucasians: no association with schizophrenia."
RL Am. J. Med. Genet. 54:384-390(1994).
CC -!- FUNCTION: This is one of the five types (D1 to D5) of receptors
CC for dopamine. The activity of this receptor is mediated by G
CC proteins which inhibit adenylyl cyclase.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- POLYMORPHISM: The number of repeats is highly polymorphic and
CC varies among different alleles. The sequence shown is that of
CC allele D4.7.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L12398; AAB59386.1; -
CC EMBL; L12397; -; NOT_ANNOTATED_CDS.
CC FIC; S15079; DTHUD4.
CC Genew; HGNC:3025; DRD4.
CC MIM; 126452; -
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004952; F:dopamine receptor activity; TAS.

```


CC Note=No experimental confirmation available;
 CC Name=4;
 CC IsoId=Q9Y286-4; Sequence=VSP_002557, VSP_002558;
 CC Note=No experimental confirmation available;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Predominantly expressed by resting and
 CC activated natural killer cells and at lower levels by granulocytes
 CC and monocytes. High expression found in placenta, liver, lung,
 CC spleen, and peripheral blood leukocytes.
 CC -!- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred
 CC to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).
 CC This motif is involved in downmodulation of cellular responses.
 CC The phosphorylated ITIM motif binds to the SH2 domain of
 CC PTN2/SHP-1.
 CC -!- PTM: Tyrosine phosphorylated.
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
 CC (sialic acid binding Ig-like lectin) family.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
 CC frameshift in position 406.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF170485; AAF12759.1; -
 CC EMBL; AJ007395; CAB46011.1; -
 CC EMBL; AJ130710; CAB51126.1; -
 CC EMBL; AJ130711; CAB51127.1; -
 CC EMBL; AJ130712; CAB51128.1; -
 CC EMBL; AJ130713; CAB51129.1; -
 CC EMBL; AJ178981; AAF44346.1; ALT_FRAME.
 CC EMBL; AF193441; AAF06790.1; -
 CC GenBank; HGNC:10876; SIGLEC7.
 CC MIN; 50410; -
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0005330; F:lectin; TAS.
 CC GO; GO:0004872; F:receptor activity; TAS.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC Pfam; PF00047; Ig; 3.
 CC SMART; SM00408; IGC2; 1.
 CC PROSITE; PS00835; IG_LIKE; 2.
 CC Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
 CC Immunoglobulin domain; Repeat; Phosphorylation; Alternative splicing.
 CC SIGNAL 1 18
 CC CHAIN 19 467
 CC DOMAIN 19 353
 CC TRANSMEM 354 376
 CC DOMAIN 377 467
 CC DOMAIN 39 122
 CC DOMAIN 150 233
 CC DOMAIN 240 336
 CC SITE 435 440
 CC DISULFID 46 106
 CC DISULFID 168 217
 CC DISULFID 276 320
 CC CARBOHYD 105 105
 CC CARBOHYD 142 142
 CC CARBOHYD 165 165
 CC CARBOHYD 229 229
 CC CARBOHYD 235 235
 CC CARBOHYD 242 242
 CC CARBOHYD 260 260
 CC CARBOHYD 334 334
 CC VARSPLIC 145 238
 CC GTSVSPPLHPTTRSSVLTIPQPHGHTSLCQVTLPGAV
 CC TNRRTQLAVSY -> D (in isoform 2).
 CC /FTid=VSP_002555.

FT VARSPLIC 145 145 A -> E (in isoform 3).
 FT /FTid=VSP_002556.
 FT A -> G (in isoform 4).
 FT /FTid=VSP_002557.
 FT VARSPLIC 146 467 Missing (in isoform 3 and isoform 4).
 FT /FTid=VSP_002558.
 FT CONFLICT 42 42 V -> A (IN REF. 4).
 SQ SEQUENCE 467 AA; 51142 MW; 8AFE44462B001F52 CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 467;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 379 GVLLGAVG 386
 Db 350 GVLLGAVG 357
 RESULT 26
 CPBK MOUSE STANDARD; PRT; 470 AA.
 ID CPBK MOUSE
 AC Q62397;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2B20 (EC 1.14.14.1) (CYPIIIB20) (P24) (Fragment).
 GN CYP2B20.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAUB/c; TISSUE=Liver;
 RX MEDLINE=96428606; PubMed=8631708;
 RA Damon M., Fautrel A., Marc N., Guilloze A., Corcos L.;
 RT "Isolation of a new mouse cDNA clone: hybrid form of cytochrome P450
 RL Biochem. Biophys. Res. Commun. 226:900-905(1996).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE LIVER AS WELL AS IN KIDNEY,
 CC LUNG AND INTESTINE.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X99715; CAA68051.1; ALT_INIT.
 CC HSPSP; P00179; 1DT6.
 CC MGD; MGI:1202389; Cyp2b20.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008068; EP450_CYP2B.
 CC Pfam; PF00067; p450; 1.
 CC PRINTS; PRO1685; EP450ICYP2B.
 CC PRINTS; PRO0385; P450.
 CC PROSITE; PS00086; Monooxygenase; P450.
 CC Oxidoreductase; Cytochrome P450; 1.
 CC Microsome; Endoplasmic reticulum; Phosphorylation.
 CC MOD_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
 CC METAL 436 436 IRON (HEME AXIAL LIGAND).
 CC NON_TER 470 470

SQ SEQUENCE 470 AA; 53357 MW; 8B9CF3E2EA622642 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1153 LLLALLV 1160
|||||
DB 6 LLLALLV 13

RESULT 27
XYLC_PSEPU STANDARD; PRT; 487 AA.
AC P43503;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] [EC 1.2.1.28].
GN XYLC.
OS Pseudomonas putida.
OG Plasmid TOL pW53.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RP PLASMID=TOL pW53.
RX MEDLINE=95173094; PubMed=7868591;
RA Inoue J., Shaw J.P., Reik M., Harayama S.;
RT "Overlapping substrate specificities of benzaldehyde dehydrogenase
(the xylC gene product) and 2-hydroxymuconic semialdehyde
dehydrogenase (the xylG gene product) encoded by TOL plasmid pW50 of
Pseudomonas putida.";
RL J. Bacteriol. 177:1196-1201 (1995).
RN [2]
RN SEQUENCE OF 1-53.
RP PLASMID=TOL pW53;
RX MEDLINE=9113163; PubMed=1989592;
RA Chalmers R.M., Keen J.N., Fawson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
dehydrogenases from the benzyl alcohol and mandelate pathways in
Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
acid compositions and immunological cross-reactions.";
RL Biochem. J. 273:99-107 (1991).
RN [3]
RN CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate +
NADH.
CC -1- SUBUNIT: Homotetramer.
CC -1- MISCELLANEOUS: Optimal pH is 9.0.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U15151; AAA66218.1; --
DR EMBL; D63341; BAA09661.1; --
DR PIR; T47107; T47107.
DR HSSP; P51977; 1EXS.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR CYS; 1.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD; Plasmid.
FT NP BIND 232 237 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 254 254 BY SIMILARITY.
FT ACT_SITE 288 288 BY SIMILARITY.
SQ SEQUENCE 487 AA; 51897 MW; 093CB3E9487AF384 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 626 ALGNAVIL 633
|||||
DB 171 ALGNAVIL 178

RESULT 28
CPBI_RAT STANDARD; PRT; 491 AA.
AC P00176;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B1 (EC 1.14.14.1) (CYPIB1) (P450-B) (P450-PB1 and
P450-PB2) (P450-LM2).
GN CYP2B1 OR CYP2B-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE OF 6-491 FROM N.A. (ISOZYMES PB1 AND PB2).
RP MEDLINE=8222224; PubMed=6953431;
RX Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RT "Primary structure of a cytochrome P-450: coding nucleotide sequence
of phenobarbital-inducible cytochrome P-450 cDNA from rat liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2793-2797 (1982).
RN [2]
RN REVISIONS TO 166; 292 AND 378 (ISOZYMES PB1 AND PB2).
RP Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RL Proc. Natl. Acad. Sci. U.S.A. 79:5443-5443 (1982).
RN [3]
RN SEQUENCE OF 1-22.
RX MEDLINE=7919411; PubMed=109438;
RA Botelho L.H., Ryan D.E., Levin W.;
RT "Amino acid compositions and partial amino acid sequences of three
highly purified forms of liver microsomal cytochrome P-450 from rats
treated with polychlorinated biphenyls, phenobarbital, or 3-
methylcholanthrene.";
RL J. Biol. Chem. 254:5635-5640 (1979).
RN [4]
RN PHOSPHORYLATION.
RP MEDLINE=90059885; PubMed=2583091;
RA Pyerin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome
P-450.";
RL EMBL J. 8:3003-3010 (1989).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: By phenobarbital.
CC -1- PTM: Phosphorylation is accompanied by a decrease in enzyme
activity.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; J00719; AAA1024.1; --

```

DR EMBL; M37134; AAC42028.1; -
DR PIR; A00176; O4RTPB.
DR PIR; A21162; O4RTP2.
DR HSP; P00179; IDT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450CYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
FT METAL 436 436 IRON (HEME AXIAL LIGAND).
FT VARIANT 303 303 S -> G (IN ISOZYME PB2).
FT VARIANT 321 322 AE -> TV (IN ISOZYME PB2).
FT VARIANT 337 337 L -> P (IN ISOZYME PB2).
FT VARIANT 339 339 T -> S (IN ISOZYME PB2).
FT VARIANT 344 344 S -> T (IN ISOZYME PB2).
SQ SEQUENCE 491 AA; 55933 MW; 74615501AD5497DD CRC64;

Query Match 0.78; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLLV 1160
| | | | |
Db 6 LLLALLLV 13

RESULT 29
CPB2 RAT STANDARD; PRT; 491 AA.
AC P04157; O64582;
DT 20-MAR-1987 (Rel. 04, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B2 (BC 1.14.14.1) (CYP1B2) (P450E) (P450 PB4).
GN CYP2B2 OR CYP2B-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=83247397; PubMed=6306654;
RA Mizukami Y., Sogawa K., Suwa Y., Muramatsu M., Fujii-Kuriyama Y.;
RT "Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver."
RL Proc. Natl. Acad. Sci. U.S.A. 80:3958-3962(1983).
RN [2]
SEQUENCE.
RP MEDLINE=86059379; PubMed=3877725;
RA Frey A.B., Waxman D.J., Kreibich G.;
RT "The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4. Production and characterization of site-specific antibodies."
RL J. Biol. Chem. 260:15253-15265(1985).
RN [3]
SEQUENCE OF 168-491 FROM N.A.
RP MEDLINE=84159487; PubMed=6689485;
RA Phillips I.R., Shephard E.A., Ashworth A., Rabin B.R.;
RT "Cloning and sequence analysis of a rat liver cDNA coding for a phenobarbital-inducible microheterogenous cytochrome P-450 variant: regulation of its messenger level by xenobiotics."
RL Gene 26:41-52(1983).
RN [4]
SEQUENCE OF 281-491 FROM N.A.
RP MEDLINE=83291091; PubMed=6688421;
RA Kumar A., Raphael C., Adesnik M.;
RT "Cloned cytochrome P-450 cDNA. Nucleotide sequence and homology to multiple phenobarbital-induced mRNA species."
RL J. Biol. Chem. 258:11280-11284(1983).

[5]
RR ERRATUM.
RA Kumar A., Raphael C., Adesnik M.;
RL J. Biol. Chem. 259:6039-6039(1984).
[6]
RN SEQUENCE OF 323-431 FROM N.A.
RX MEDLINE=86205943; PubMed=3458196;
RA Atchison M.L., Adesnik M.;
RT "Gene conversion in a cytochrome P-450 gene family."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2300-2304(1986).
[7]
RN SEQUENCE OF 385-491 FROM N.A.
RX MEDLINE=84153837; PubMed=6322758;
RA Affolter M., Anderson A.;
RT "Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450e and P-450b cDNAs and cytochrome P-450e-like genes."
RL Biochem. Biophys. Res. Commun. 118:655-662(1984).
[8]
RN SEQUENCE OF 1-20 FROM N.A.
RX MEDLINE=88273074; PubMed=2839467;
RA Hashimoto T., Matsumoto T., Nishizawa M., Kawabata S.,
RA Morohashi K., Handa S., Omura T.;
RT "A mutant rat strain deficient in induction of a phenobarbital-inducible form of cytochrome P-450 in liver microsomes."
RL J. Biochem. 103:487-492(1988).
[9]
RN PHOSPHORYLATION.
RP MEDLINE=90059885; PubMed=2583091;
RA Pyerin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome P-450."
RL EMBO J. 8:3003-3010(1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By phenobarbital.
CC -!- PTM: Phosphorylation is accompanied by a decrease in enzyme activity.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; J00728; AAA41056.1; -
EMBL; J00720; AAA41056.1; JOINED.
EMBL; J00721; AAA41056.1; JOINED.
EMBL; J00722; AAA41056.1; JOINED.
EMBL; J00723; AAA41056.1; JOINED.
EMBL; J00724; AAA41056.1; JOINED.
EMBL; J00725; AAA41056.1; JOINED.
EMBL; J00726; AAA41056.1; JOINED.
EMBL; K00596; AAA41029.1; -
EMBL; K01626; AAA41037.1; -
EMBL; K01721; AAA41026.1; -
EMBL; D00250; BAA00181.1; -
EMBL; M13234; AAA41057.1; -
PIR; A21162; O4RTP2.
PIR; A60822; A60822.
HSP; P00179; IDT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.

```

Pfam; PF00067; P450; 1.
 DR PRINTS; PRO1685; EP450ICYP2B.
 DR PRINTS; PRO0385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Phosphorylation.
 FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
 FT METAL 436 436 IRON (HEME AXIAL LIGAND).
 FT CONFLICT 292 292 L -> P (IN REF. 2).
 FT CONFLICT 321 321 T -> A (IN REF. 2 AND 4).
 FT CONFLICT 322 322 E -> V (IN REF. 1).
 FT CONFLICT 438 438 G -> D (IN REF. 4).
 FT CONFLICT 444 444 N -> K (IN REF. 3).
 FT CONFLICT 473 473 K -> M (IN REF. 1).
 FT CONFLICT 476 476 G -> D (IN REF. 2 AND 4).
 SQ SEQUENCE 491 AA; 55932 MW; 00CB6937FDD44BC CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALLIV 1160
 DB 6 LLLALLIV 13
 RESULT 30
 CPBA_MOUSE STANDARD; PRT; 500 AA.
 AC P12791;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2B10 (EC 1.14.14.1) (CYP11B1) (Testosterone 16-alpha
 DE hydroxylase) (P450-16-alpha) (Clone PF3/46).
 GN CYP2B10 OR CYP2B-10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=69118235; PubMed=3219345;
 RX Noshiro M., Iakso M., Kawajiri K., Negishi M.;
 RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)
 RT of testosterone 16 alpha-hydroxylase in mouse liver, chromosome
 RT localization, and cloning of p-450 cDNA.";
 RL Biochemistry 27:6434-6443 (1988).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 EMBL; M21856; AAA0425.1; -;
 DR PIR; B31047; B31047.
 DR HSSP; P00179; 1DT6.
 DR MGD; MGI:88538; Cyp2b10.
 DR InterPro; IPR001128; Cytochrome P450.
 DR IntronPro; IPR008068; EP450_CYP2B.
 PFam; PF00067; P450; 1.

DR PRINTS; PRO1685; EP450ICYP2B.
 DR PRINTS; PRO0385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Phosphorylation.
 FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
 FT METAL 445 445 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 500 AA; 56743 MW; F660A008D0FBA94 CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 500;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALLIV 1160
 DB 6 LLLALLIV 13
 RESULT 31
 VE1_BP2V STANDARD; PRT; 604 AA.
 AC P11298;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Replication protein E1.
 GN E1.
 OS Bovine papillomavirus type 2.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Groff D.E., Mitra R., Lancaster W.D.;
 RL Submitted (MAY-1989) to the EMBL/GenBank/DBSJ databases.
 CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
 CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
 CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
 CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 EMBL; M20219; AAA66833.1; -;
 DR PIR; C31169; W1WLB2.
 DR InterPro; IPR001177; Papillom_E1.
 DR Pfam; PF00519; E1; 1.
 DR Pfam; PF00524; E1_N; 1.
 KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
 KW Nuclear protein.
 FT NP_BIND 432 439 ATP (POTENTIAL).
 SQ SEQUENCE 604 AA; 68077 MW; D2D7036ADE88A9DD CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 604;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 927 LAAGSDSN 934
 DB 336 LAAGSDSN 343
 RESULT 32
 VE1_BP2V STANDARD; PRT; 605 AA.
 ID VE1_BP2V
 AC P03116; Q9WVH1;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1996 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DN Replication protein E1.
 GN E1
 OS Bovine papillomavirus type 1.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10559;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=63012974; PubMed=6289124;
 RA Chen E.Y., Howley P.M., Levinson A.D., Seeburg P.H.;
 RT "The primary structure and genetic organization of the bovine
 papillomavirus type 1 genome.";
 RL Nature 299:529-534(1982).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91122053; PubMed=1846806;
 RA Ustav M., Stenlung A.;
 RT "transient replication of BPV-1 requires two viral polypeptides
 encoded by the E1 and E2 open reading frames.";
 RL EMBO J. 10:449-457(1991).
 RN (3)
 RP CHARACTERIZATION.
 RX MEDLINE=93281701; PubMed=9389467;
 RA Yang L., Mohr I., Fouts E., Lim D.A., Nohaile M., Botchan M.;
 RT "The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA
 helicase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5086-5090(1993).
 CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
 VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
 PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
 CONTAINS BINDING SITES FOR BOTH PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X02346; CAB46511.1; -;
 DR PIR; A03663; WIMLEB.
 DR PDB; 1F08; 1F-WAY-01.
 DR TRANSFAC; T020203; -;
 DR InterPro; IPR001177; Papillom_E1.
 DR Pfam; PF00519; E1; 1.
 DR Pfam; PF00524; E1 N; 1.
 KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
 KW Nuclear protein; 3D-structure.
 FT NP BIND 433 440 ATP (POTENTIAL)
 FT SEQUENCE 605 AA; 65190 MW; C6400B7E8F605E0B CRC64;
 SQ
 Query Match 0.7%; Score 8; DB 1; Length 605;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 927 LAAGSDSN 934
 Db 337 LAAGSDSN 344
 |||||
 RESULT 33
 TMS6_HUMAN STANDARD; PRT; 811 AA.
 ID TMS6_HUMAN
 AC Q8IU80; Q8IU82; Q8IXV8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Transmembrane protease, serine 6 (BC 3.4.21.-) (Matrptase-2).
 GN TMPRSS6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dutham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehara H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon I., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudo H.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Z., Wang Z., White J., Willingham D., Wu H., Yac Z.,
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,
 RA Scheet P., Walker C., Waisley A., Wohlmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L.W., Marais E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edlmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Serousai E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tlahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Brain;
 MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

FT				(POTENTIAL).
FT	DOMAIN	77	811	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	213	336	CUB 1.
FT	DOMAIN	335	452	CUB 2.
FT	DOMAIN	457	489	LDL-RECEPTOR CLASS A 1.
FT	DOMAIN	490	526	LDL-RECEPTOR CLASS A 2.
FT	DOMAIN	530	567	LDL-RECEPTOR CLASS A 3.
FT	DOMAIN	577	811	SERINE PROTEASE.
FT	ACT_SITE	617		CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	668		CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	762	762	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	216	216	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	433	433	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	453	453	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	518	518	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC	409	461	LCGLRLQYAKRIPVAVTACITINFSQISLTGPGVRVHY GLYNQSDPCPE -> YHFLSGLMLPLPPPSLPSSVTVP SLAEQVNLRGAARGSGWGCQACP (in isoform 2).
FT				/FTid=VSP_008379. Missing (In isoform 2).
FT	VARSPLIC	462	811	/FTid=VSP_008380. A -> V (IN REF. 4).
FT	CONFLICT	116	116	A -> V (IN REF. 4).
FT	SEQUENCE	811 AA;	89999 MW;	7EEF193F655DE9D CRC64;
SQ				
	Query Match	0.7%;	Score 8;	DB 1; Length 811;
	Best Local Similarity	100.0%;	Pred. No. 33;	Mismatches 0; Indels 0; Gaps 0;
	Matches	8;	Conservative	0;
QY		1155 LLALLVLA 1162		
Db		59 LLALLVLA 66		
RESULT 34				
ID	ITAS_XENIA			
ID	ITAS_XENLA	STANDARD;	PRT;	1050 AA.
AC	Q06274;			
DT	15-JUL-1998	(Rel. 36, Created)		
DT	15-JUL-1998	(Rel. 36, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Integrin alpha-5 precursor (fibronectin receptor alpha subunit)			
DE	(integrin alpha-F) (VLA-5).			
OS	Xenopus laevis	(African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=95344994; PubMed=7619730;			
RA	Jos T.O., Whittaker C.A., Meng F., Desimone D.W., Gnau V.,			
RA	Hausen P.;			
RT	"Integrin alpha 5 during early development of Xenopus laevis.";			
RL	Mech. Dev. 50:187-199(1995).			
RL	[2]			
RP	SEQUENCE OF 318-393 FROM N.A.			
RX	MEDLINE=94008528; PubMed=8404528;			
RA	Whittaker C.A., Desimone D.W.;			
RT	"Integrin alpha subunit mRNAs are differentially expressed in early			
RL	Xenopus embryos."			
RL	Development 117:1239-1249(1993).			
CC	-I- FUNCTION: INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN.			
CC	-I- SUBUNIT: THE SEQUENCE R-G-D IN ITS LIGANDS.			
CC	-I- SUBUNIT: HETEROIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA			
CC	SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A			
CC	DISULFIDE BOND. ALPHA-5 ASSOCIATES WITH BETA-1.			
CC	-I- SURCELLULAR LOCATION: Type I membrane protein.			
CC	-I- SIMILARITY: Belongs to the integrin alpha chain family.			
CC	-I- SIMILARITY: Contains 7 FG-GAP repeats.			

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: U16883; AAA99668.1; -;
 EMBL: L10191; AAA16249.1; -;
 PIR: I51527; I51527.
 HSSP: P06756; 1JUV2.
 InterPro: IPR000413; Integrin_alpha.
 Pfam: PF01839; FG-GAP; 4.
 Pfam: PF00357; integrin_A; 1.
 SMART: SMO0191; int_alpha; 5.
 PROSITE: PS00242; INTEGRIN ALPHA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Repeat.
 SIGNAL 1 32 POTENTIAL.
 CHAIN 33 1050 INTEGRIN ALPHA-5.
 CHAIN 33 932 INTEGRIN ALPHA-5 HEAVY CHAIN (POTENTIAL).
 CHAIN 933 1050 INTEGRIN ALPHA-5 LIGHT CHAIN (POTENTIAL).
 DOMAIN 33 996 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 997 1022 POTENTIAL.
 DOMAIN 1023 1050 CYTOPLASMIC (POTENTIAL).
 REPEAT 48 110 FG-GAP 1.
 REPEAT 120 188 FG-GAP 2.
 REPEAT 189 246 FG-GAP 3.
 REPEAT 259 312 FG-GAP 4.
 REPEAT 313 373 FG-GAP 5.
 REPEAT 379 438 FG-GAP 6.
 REPEAT 442 494 FG-GAP 7.
 CA_BIND 324 332 POTENTIAL.
 CA_BIND 390 398 POTENTIAL.
 CA_BIND 454 462 POTENTIAL.
 SITE 1025 1029 GFFKR MOTIF.
 DISULFID 90 99 BY SIMILARITY.
 DISULFID 145 166 BY SIMILARITY.
 DISULFID 182 195 BY SIMILARITY.
 DISULFID 502 513 BY SIMILARITY.
 DISULFID 519 575 BY SIMILARITY.
 DISULFID 636 642 BY SIMILARITY.
 DISULFID 708 721 BY SIMILARITY.
 DISULFID 862 910 INTERCHAIN (BY SIMILARITY).
 CARBOHYD 917 922 BY SIMILARITY.
 CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 507 507 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 521 521 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 600 600 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 861 861 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 1050 AA; 115961 MW; 10ED96153538D918 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1050;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 RFGSSIAS 548
 |||||

Db 380 RFGSSIAS 387
 |||||

RESULT 35
 ITAL DROME
 ID ITAL DROME STANDARD; PRT; 1146 AA.
 AC Q24247; Q5VYF6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-PS1 precursor (Position-specific antigen 1, alpha chain) (Protein multiple edematous wings).
 GN MEW OR CGI771.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=Oregon-R;
 RX MEHLIN=94059764; PubMed=8240969;
 RA Wehrli M., Diantonio A., Pearnley I.M., Smith R.J., Wilcox M.;
 RT "Cloning and characterization of alpha PSI, a novel Drosophila melanogaster integrin.";
 RL Mech. Dev. 43:21-36(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dursin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: INTEGRIN ALPHA-PS1/BETA-PS IS A RECEPTOR FOR LAMININ.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
 CC DISULFIDE BOND. ALPHA-PS1 ASSOCIATES WITH BETA-PS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous

CC
CC Gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL curation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.eib-sib.ch/announce/>
CC or send an email to license@sib.ch).
CC

[illegible]

```

Query Match      0.7%; Score 8; DB 1; Length 1146;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1151  GGLLLLLL 1158
          |||||
          |||||
DB       1094  GGLLLLLL 1101
          |||||
          |||||

```

RESULT 36	ITAM_HUMAN	STANDARD;	PRT; 1152 AA.
ID	ITAM_HUMAN		
AC	P11215;		
CD	01-JUL-1989	(Rel. 11, Created)	
DDT	01-OCT-1996	(Rel. 34, Last sequence update)	
DDT	10-OCT-2003	(Rel. 42, Last annotation update)	
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha		

subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor M01)
DE (Neutrophil adherence receptor).
DE ITGAM OR CR3A OR CD11B.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=98315033; PubMed=2457584;
CORBI A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
ERRA "the human leukocyte adhesion glycoprotein Mac-1 (complement receptor
type 3, CD11b) alpha subunit. Cloning, primary structure, and
ERRT relation to the integrins, von Willebrand factor and factor B.";
ERRT J. Biol. Chem. 263:12403-12411(1988).
RN [2]
SEQUENCE FROM N.A.
MEDLINE=9839051; PubMed=2833753;
ERRA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
ERRA "Molecular cloning of the alpha subunit of human and guinea pig
ERRT leukocyte adhesion glycoprotein M01: chromosomal localization and
ERRT homology to the alpha subunits of integrins.";
ERRT Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
SEQUENCE FROM N.A.
MEDLINE=88257215; PubMed=2454931;
ERRA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
ERRA "Amino acid sequence of the alpha subunit of human leukocyte adhesion
ERRT receptor M01 (complement receptor type 3).";
ERRT J. Cell Biol. 106:2153-2158(1988).
RN [4]
SEQUENCE FROM N.A.
MEDLINE=93123748; PubMed=8419480;
ERRA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
ERRA "Structural analysis of the CD11b gene and phylogenetic analysis of
ERRT the alpha-integrin gene family demonstrate remarkable conservation of
ERRT genomic organization and suggest early diversification during
ERRT evolution.";
ERRT J. Immunol. 150:480-490(1993).
RN [5]
SEQUENCE OF 9-1153 FROM N.A.
MEDLINE=8908893; PubMed=2563162;
ERRA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
ERRA Roth G.J.;
ERRA "cDNA sequence for the alpha M subunit of the human neutrophil
ERRT adherence receptor indicates homology to integrin alpha subunits.";
ERRT Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN [6]
SEQUENCE OF 1-9 FROM N.A.
MEDLINE=92073318; PubMed=1683702;
ERRA Shelley C.S., Arnaout M.A.;
ERRA "the promoter of the CD11b gene directs myeloid-specific and
ERRT developmentally regulated expression.";
ERRT Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
SEQUENCE OF 1-9 FROM N.A.
ERRA TISSUE=Blood;
ERRA MEDLINE=92144986; PubMed=1346576;
ERRA Pahl H.L., Rosmarin A.G., Tenen D.G.;
ERRA "Characterization of the myeloid-specific CD11b promoter.";
ERRT Blood 79:865-870(1992).
RN [8]
SEQUENCE OF 17-31.
MEDLINE=87076671; PubMed=3539202;
ERRA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
ERRA "N-terminal sequence of human leukocyte glycoprotein M01:
ERRT conservation across species and homology to platelet IIb/IIIa.";
ERRT Biochim. Biophys. Acta 874:369-371(1986).
RN [9]
X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
ERRA MEDLINE=95717458; PubMed=7867070;
ERRA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
ERRA "Crystal structure of the A domain from the alpha subunit of integrin
ERRT

RT CR3 (CD11b/CD18).";
 RL Cell 80:631-638(1995).
 [10]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
 RX MEDLINE=96363671; PubMed=8747460;
 RA Lee J.C., Bankston L.A., Annot M.A., Liddington R.C.;
 RT "Two conformations of the integrin A-domain (I-domain): a pathway for
 RT activation?";
 RL Structure 3:1333-1340(1995).
 [11]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
 RX MEDLINE=98362595; PubMed=9687375;
 RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
 RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
 RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
 RA Mutchler V.T., Tomich C.S., Waterpaugh K.D., Wiley V.H.;
 RT "Cation binding to the integrin CD11b I domain and activation model
 RT assessment.";
 RL Structure 6:923-935(1998).
 [12]
 RP 2D-STRUCTURE MODELING OF 17-616.
 RX MEDLINE=9826734; PubMed=9560195;
 RA Oxvig C., Springer T.A.;
 RT "Experimental support for a beta-propeller domain in integrin alpha-
 RT subunits and a calcium binding site on its lower surface.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
 CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
 CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
 CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
 CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
 CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
 CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
 CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
 CC OF FIBRINOGEN GAMMA CHAIN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 CC GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC -----
 DR EMBL; J03925; AAA59544.1; -;
 DR EMBL; M18044; AAA59491.1; -;
 DR EMBL; J04145; AAA59903.1; -;
 DR EMBL; S52227; AAB24821.1; -;
 DR EMBL; S52152; AAB24821.1; JOINED.
 DR EMBL; S52133; AAB24821.1; JOINED.
 DR EMBL; S52154; AAB24821.1; JOINED.
 DR EMBL; S52155; AAB24821.1; JOINED.
 DR EMBL; S52157; AAB24821.1; JOINED.
 DR EMBL; S52159; AAB24821.1; JOINED.
 DR EMBL; S52161; AAB24821.1; JOINED.
 DR EMBL; S52164; AAB24821.1; JOINED.
 DR EMBL; S52165; AAB24821.1; JOINED.
 DR EMBL; S52167; AAB24821.1; JOINED.
 DR EMBL; S52169; AAB24821.1; JOINED.
 DR EMBL; S52170; AAB24821.1; JOINED.
 DR EMBL; S52173; AAB24821.1; JOINED.
 DR EMBL; S52174; AAB24821.1; JOINED.

DR EMBL; S52180; AAB24821.1; JOINED.
 DR EMBL; S52181; AAB24821.1; JOINED.
 DR EMBL; S52184; AAB24821.1; JOINED.
 DR EMBL; S52189; AAB24821.1; JOINED.
 DR EMBL; S52191; AAB24821.1; JOINED.
 DR EMBL; S52192; AAB24821.1; JOINED.
 DR EMBL; S52203; AAB24821.1; JOINED.
 DR EMBL; S52212; AAB24821.1; JOINED.
 DR EMBL; S52213; AAB24821.1; JOINED.
 DR EMBL; S52216; AAB24821.1; JOINED.
 DR EMBL; S52219; AAB24821.1; JOINED.
 DR EMBL; S52220; AAB24821.1; JOINED.
 DR EMBL; S52221; AAB24821.1; JOINED.
 DR EMBL; S52222; AAB24821.1; JOINED.
 DR EMBL; S52226; AAB24821.1; JOINED.
 DR EMBL; M76724; AAA58410.1; -;
 DR EMBL; M84477; AAA51960.1; -;
 DR PIR; A31108; RWU1B.
 DR PDB; 1A8X; 17-JUN-98.
 DR PDB; 1BHO; 18-NOV-98.
 DR PDB; 1BHO; 18-NOV-98.
 DR PDB; 1IDN; 25-NOV-98.
 DR PDB; 1IDO; 01-AUG-96.
 DR PDB; 1JLM; 11-JAN-97.
 DR PDB; 1MIU; 07-AUG-02.
 DR GENE; HGNC:6149; ITGAM.
 DR MIM; 120980; -;
 DR GO; GO:0008305; C:cell adhesion receptor activity; TAS.
 DR GO; GO:0004895; F:cell adhesion; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
 FT SIGNAL 1 16
 FT CHAIN 17 1152 INTEGRIN ALPHA-M.
 FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1105 1128 POTENTIAL.
 FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
 Query Match 0.7%; Score 8; DB 1; Length 1152;
 Best Local Similarity 100.0%; Pred.No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1151 GGLLLAL 1158
 DB 1115 GGLLLAL 1122
 RESULT 37
 DD37 HUMAN
 ID DD37 HUMAN STANDARD; PRT; 1157 AA.
 AC Q81Y37; Q85U17; Q9P211;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE Probable ATP-dependent helicase DHX37 (DEAH-box protein 37).
 GN DHX37 OR DD37 OR KIAA1517.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_taxid=9606;
 RN [1]

RC TISSUE=Brain, and Duodenum;
RX MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE OF 178-1157 FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro."
DNA Res. 7:143-150(2000).
RL
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAD
subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; BC002575; AA02575.1; ALT_INIT.
DR EMBL; BC037964; AA037964.1; -.
DR EMBL; AB040950; BAA96041.1; -.
DR Genbank; HGNC:17210; DHX37.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF04408; HA2; 1.
DR SMART; SM00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC_C; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; FALSE_NEG.
KW Hydrolase; Helicase; ATP-binding.
FT NP BIND 275 282 ATP (POTENTIAL).
FT SITE 372 375 DEAD BOX.
FT S -> G (IN REF. 1; AA02575).
FT A -> S (IN REF. 2).
FT R -> Q (IN REF. 1; AA02575).
FT YLLAEYCEWLFQAMHPDIEKAWPTTHV -> CEFDQOGGV
FT GVDNRGSLRQGLCACTVSPGLAGSGTAAQLFAT (IN
FT REF. 2).
FT SEQUENCE 1157 AA; 129544 MW; 49332175221B30C5 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1157;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 380 VLLGAVGA 387
| | | | |
DB 857 VLLGAVGA 864

RESULT 38
ITAD_HUMAN
ID ITAD_HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.I., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
3.";
J. Biol. Chem. 275:8959-8969(2000).
RL [3]
RN SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RP MEDLINE=96257236; PubMed=8666289;
RX Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
beta 2-integrin alpha subunit.";
Gene 171:291-294(1996).
RL [4]
RN INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "alpha2beta2 integrin is expressed on human eosinophils and functions
as an alternative ligand for vascular cell adhesion molecule 1
(VCAM-1)."
J. Exp. Med. 188:2187-2191(1998).
RL [5]
RN INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
binding interface between I domain and VCAM-1.";
J. Immunol. 163:1984-1990(1999).
RL -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
BORNE PATHOGENS, PARTICULATE MATTER, AND SENESENT ERYTHROCYTES
FROM THE BLOOD.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```
KW EGF-like domain; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1173
FT DOMAIN 23 235
FT DOMAIN 23 235
FT DOMAIN 23 224
FT DOMAIN 319 376
FT DOMAIN 382 432
FT DOMAIN 438 493
FT DOMAIN 495 550
FT DOMAIN 550 590
FT DOMAIN 591 648
FT DOMAIN 649 693
FT DOMAIN 726 761
FT DOMAIN 762 784
FT DOMAIN 785 820
FT DOMAIN 821 843
FT DOMAIN 844 881
FT DOMAIN 882 917
FT DOMAIN 918 953
FT DOMAIN 954 1173
FT SITE 929 931
FT SITE 394 426
FT DISULFID 398 431
FT DISULFID 450 487
FT DISULFID 454 492
FT DISULFID 465 477
FT DISULFID 507 544
FT DISULFID 511 549
FT DISULFID 522 534
FT DISULFID 554 565
FT DISULFID 559 575
FT DISULFID 578 589
FT DISULFID 595 611
FT DISULFID 602 620
FT DISULFID 623 647
FT DISULFID 653 666
FT DISULFID 660 679
FT DISULFID 681 692
FT DISULFID 708 716
FT DISULFID 721 741
FT DISULFID 757 777
FT DISULFID 780 800
FT DISULFID 816 836
FT DISULFID 839 859
FT DISULFID 877 897
FT DISULFID 913 933
FT DISULFID 949 1170
FT CARBOHYD 155 155
FT CARBOHYD 158 158
FT CARBOHYD 250 250
FT CARBOHYD 363 363
FT CARBOHYD 705 705
FT CARBOHYD 711 711
FT CARBOHYD 1070 1070
SQ SEQUENCE 1173 AA; 130019 MW; A9F036D6516C0F24 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1173;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 486 SVDIDGDG 493
Db 801 SVDIDGDG 808

RESULT 41
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570: Q9NZU9.
DC 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
```

Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (integrin alpha-IEL).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.

RC TISSUE=Leukemia, and Lymphocytes;

RX MEDLINE=94164962; PubMed=8119947;

RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;

RA "Molecular Cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";

RL J. Biol. Chem. 269:6016-6025(1994).

RN [2]

RP REVISIONS TO 88-114.

RA Parker C.M.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 53-1179 FROM N.A.

RC TISSUE=Fetal kidney;

RX MEDLINE=20138496; PubMed=10673275;

RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shotelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;

RA "The genomic region encompassing the nephropathic cystinosis gene (CNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";

RL Genome Res. 10:165-173(2000).

RN [4]

RP MUTAGENESIS OF ASP-109 AND PHE-316.

RX MEDLINE=20400502; PubMed=10837471;

RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;

RA "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";

RL J. Biol. Chem. 275:25652-25664(2000).

CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL EPITHELIAL CELLS.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; L25851; AAB59359.2; -.

DR EMBL; AF168787; AAF43107.1; -.

DR PIR; A53213; A53213.

DR HSP; P11215; IABX.

DR Genew; HGNC:6147; ITGAE.

DR MIM; 604682; -.

DR GO; GO:0008305; C:integrin complex; TAS.

RESULT 42
PMPC CHLMU STANDARD; PRT; 1460 AA.
ID Q9DJY1; (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmc precursor (Polymorphic membrane protein C).
DE PMPC OR TC0695.
GN Chlamydia muridarum.
OS Chlamydiales; Chlamydiales; Chlamydia.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / N199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AE002338; AAF39511.1; -.
DR PIR; D81675; D81675.
DR TIGR; TC0695; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; Chlamydia_PMP; 2.
DR TIGRfam; TIGR01414; autotrans_bar1; 1.
DR TIGRfam; TIGR01376; POMP repeat; 8.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1460 PROBABLE OUTER MEMBRANE PROTEIN PMPC.
SQ SEQUENCE 1460 AA; 154781 MW; C1F033E7907AC3AC CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1460;
Best Local Similarity 100.0%; Pred.No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 583 SILKTPKQ 590
DB 1311 SILKTPKQ 1318
RESULT 43
SET2-CABEL STANDARD; PRT; 1507 AA.
ID SET2-CABEL
AC Q18221; Q9506; Q9507;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Protein set-2
GN SET-2 OR C36E6.9.
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]

InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWF_A.
Pfam; PF01839; FG-GAP_3.
Pfam; PF00357; integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int alpha; 3.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS00234; VWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Polymorphism; Magnesium; Calcium.
FT SIGNAL 1 18
FT CHAIN 19 1179 INTEGRIN ALPHA-E.
FT CHAIN 19 1179 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 173 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1125 1147 POTENTIAL.
FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 181 198 GLU-RICH (ACIDIC).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 145 199 VWA.
FT DOMAIN 200 391 FG-GAP 3.
FT REPEAT 401 456 FG-GAP 4.
FT REPEAT 457 506 FG-GAP 5.
FT REPEAT 510 571 FG-GAP 6.
FT REPEAT 573 638 FG-GAP 7.
FT REPEAT 641 693 POTENTIAL.
FT CA_BIND 522 530 POTENTIAL.
FT CA_BIND 586 594 POTENTIAL.
FT CA_BIND 654 662 GTFKR MOTIF.
FT SITE 1150 1154 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 126 159 BY SIMILARITY.
FT DISULFID 706 762 BY SIMILARITY.
FT DISULFID 823 829 BY SIMILARITY.
FT DISULFID 893 907 BY SIMILARITY.
FT DISULFID 1008 1033 BY SIMILARITY.
FT DISULFID 1041 1057 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 360 360 D -> E.
FT VARIANT 1041 1041 /FTID=VAR_008884.
FT MUTAGEN 208 208 /FTID=VAR_008885.
FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.
FT CONFLICT 477 477 E->A: LOSS OF E-CADHERIN BINDING.
FT CONFLICT 482 482 V -> I (IN REF. 3).
FT CONFLICT 482 482 Q -> R (IN REF. 3).
FT CONFLICT 950 950 R -> W (IN REF. 3).
FT CONFLICT 1019 1019 A -> V (IN REF. 3).
SQ SEQUENCE 1179 AA; 130088 MW; E558902ED9D95E1 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1179;
Best Local Similarity 100.0%; Pred.No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 380 VLLGAVGA 387
DB 416 VLLGAVGA 423


```
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_A_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1.
DR Pfam; PF00623; RNA_pol_Rpb1_2.
DR Pfam; PF04983; RNA_pol_Rpb1_3.
DR Pfam; PF05000; RNA_pol_Rpb1_4.
DR Pfam; PF04998; RNA_pol_Rpb1_5.
DR SMART; SM00563; RPOA_N; 1.
DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN_FING 62 78 C2H2-TYPE (POTENTIAL).
FT CONFLICT 158 158 N->T (IN REF. 1).
SQ SEQUENCE 1664 AA; 186431 MW; Df65A7AA4595E6D CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1664;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 GSEITSVD 488
Db 8 GSEITSVD 15

RESULT 45
RRPL_BUNYW STANDARD; PRT; 2238 AA.
AC RPL_BUNYW
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
GN L.
OS Bunyavirus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=35304;
RN [1]
RP SEQUENCE FROM N.A.; PubMed=2596023;
RX MEDLINE=90085791;
RA Elliott R.N.;
RT "Nucleotide sequence analysis of the large (L) genomic RNA segment of
RL Bunyavirus virus, the prototype of the family Bunyaviridae."
RL Virology 173:426-436(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}[N].
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X14383; CAA32553.1;
DR PIR; A33744; RRVUBY.
DR InterPro; IPR007322; Bunya_RdRp.
DR InterPro; IPR007099; RNA_pol_NSvir.
DR Pfam; PF04196; Bunya_RdRp; 1.
KW Transferrase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2238 AA; 258668 MW; IED00AB156BAC3DA CRC64;

Query Match 0.7%; Score 8; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 FLINILKK 191
Db 1333 FLINILKK 1340
```

```
RESULT 46
LST1_MOUSE
ID LST1_MOUSE STANDARD; PRT; 95 AA.
AC O08843; O08844; Q9ZLH3, Q9ZL30;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte specific transcript 1 protein (B144 protein).
LST1 OR B144.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RN [2]
RX MEDLINE=88031493; PubMed=3117682;
RA Tsuge I., Shen F.-W., Steinmetz M., Boyse E.A.;
RT "A gene in the H-2K-D interval of the major histocompatibility
complex which is transcribed in B cells and macrophages."
RL Immunogenetics 26:378-380(1987).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX STRAIN-BALB/c; Tissue=Macrophage;
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidie U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
differential regulation, initiation, and alternative splicing."
RL Genomics 45:591-600(1997).
RN [4]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=21372017; PubMed=11478849;
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R.,
RA Weissman S.M.;
RT "Functional analysis of B144/LST1: a gene in the tumor necrosis
factor cluster that induces formation of long filopodia in eukaryotic
cells."
RL Exp. Cell Res. 268:230-244(2001).
CC -!- FUNCTION: Possible role in modulating immune responses. Has an
inhibitory effect on lymphocyte proliferation. Induces
morphological changes including production of filopodia and
microspikes when overexpressed in a variety of cell types and may
be involved in dendritic cell maturation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in
a perinuclear region corresponding to the localization of the
Golgi apparatus and throughout the cytoplasm (By
similarity).
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1; Synonyms=ml7r;
IsoId=O08843-1; Sequence=Displayed;
Name=2; Synonyms=ml2lf; Sequence=VSP_050588;
IsoId=O08843-2; Sequence=VSP_050588;
CC -!- TISSUE SPECIFICITY: Expressed in spleen and at lower levels in
thymus and liver.
CC -!- SIMILARITY: Belongs to the LST1 family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
frameshift in position 58.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M18187; AAA37273.1; ALT_FRAME.
 CC EMBL; AF000427; AAB87001.1; -.
 CC EMBL; AF000428; AAB87002.1; -.
 CC EMBL; AF109713; AAC82482.1; ALT_SEQ.
 CC PIR; I49515; I49515.
 CC MGD; MGI:1096324; Lst1.
 CC InterPro; IPR007775; LST1.
 CC Pfam; PF05083; LST1; 1.
 CC Immune response; Cell shape; Transmembrane;
 KW Alternative splicing.
 FT TRANSMEM 22 42
 FT VARSPPLIC 53 66
 FT POTENTIAL.
 FT Missing (in isoform 2).
 FT H -> D (IN REF. 2 AND 3).
 FT Y -> N (IN REF. 2 AND 3).
 FT G -> R (IN REF. 2 AND 3).
 FT Missing (IN REF. 3).
 SQ SEQUENCE 95 AA; 10325 MW; 7B22EC72CF73C2C CRC64;

Query Match 0.6%; Score 7; DB 1; Length 95;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
 DB 24 LGGLLLL 30

RESULT 47

FEL2_FELCA STANDARD; PRT; 109 AA.
 ID FEL2_FELCA
 AC P30440;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-APR-1993 (Rel. 25, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B)
 DE (Fel d 1-B) (Allergen Cat-1) (Ag4) (Fdi).
 GN CH2.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-100.
 RX MEDLINE=92052157; PubMed=1946388;
 RA Morgenstern J.P., Griffith I.J., Brauer A.W., Rogers B.L.,
 RA Bond J.F., Chapman M.D., Kuo M.-C.;
 RA "Amino acid sequence of Fel d1, the major allergen of the domestic
 cat: protein sequence analysis and cDNA cloning.";
 RT Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694(1991).
 RL [2]
 RN SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.
 RP TISSUE=Liver;
 RC MEDLINE=92241678; PubMed=1572548;
 RX Griffith I.J., Craig S., Pollock J., Yu X.-B., Morgenstern J.P.,
 RA Rogers B.L.;
 RT "Expression and genomic structure of the genes encoding Fd1, the
 RT major allergen from the domestic cat.";
 RL Gene 113:263-268(1992).
 RN [3]
 RP SEQUENCE OF 18-37, AND CHARACTERIZATION.
 RX MEDLINE=91287714; PubMed=1712088;
 RA Dufford O.A., Carreira J., Nitti G., Polo F., Lombardero M.;
 RT "Studies on the biochemical structure of the major cat allergen Felis
 RT domesticus I.";
 RL Mol. Immunol. 28:301-309(1991).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=84265679; PubMed=6747135;
 RA Leiterman K., Ohman J.L. Jr.;

RT "Cat allergen 1: Biochemical, antigenic, and allergenic properties.";
 RL J. Allergy Clin. Immunol. 74:147-153(1984).
 CC -! SUBUNIT: Heterotrimer composed of two non-covalently linked
 CC disulfide-linked heterodimer of chains 1 and 2.
 CC -! ALTERNATIVE PRODUCTS:
 CC Event=alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=1; Synonyms=CH2L;
 CC ISOID=P30440-1; Sequence=Displayed;
 CC Name=2; Synonyms=CH2S;
 CC ISOID=P30440-2; Sequence=VSP 004249;
 CC Name=3; Synonyms=CH2ST, Truncated;
 CC ISOID=P30440-3; Sequence=VSP 004248;
 CC -! TISSUE SPECIFICITY: The long form is preferentially expressed in
 CC the salivary gland, while the short form is preferentially
 CC expressed in the skin.
 CC -! ALLERGEN: Causes an allergic reaction in human. Major allergen
 CC produced by the domestic cat.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M77341; AAC41616.1; -.
 CC EMBL; X62478; CAA44345.1; -.
 CC PIR; B53283; B53283.
 CC PIR; C56413; C56413.
 CC PIR; JCL127; JCL127.
 CC InterPro; IPR006038; Uteroglobin_supf.
 KW Allergen; Glycoprotein; Signal; Polymorphism; Alternative splicing.
 FT SIGNAL 1 17
 FT CHAIN 18 109
 FT CARBOHYD 50 50
 FT VARSPPLIC 82 109
 FT
 FT VARSPPLIC 82 89
 FT
 FT VARIANT 72 72
 FT VARIANT 72 72
 FT VARIANT 74 75
 FT VARIANT 91 91
 FT VARIANT 96 96
 FT VARIANT 105 105
 FT CONFLICT 24 24
 FT CONFLICT 32 32
 SQ SEQUENCE 109 AA; 11854 MW; 857FB9CD76036CB9 CRC64;
 FTID=VSP_004248.
 TTISSSKD -> IAINFY (in isoform 2).
 /FTID=VSP_004249.
 I -> L (IN CH2LV).
 I -> V (IN CH2SV).
 RV -> KF (IN CH2SV).
 M -> T (IN CH2LV).
 Q -> E (IN CH2SV).
 N -> K (IN CH2SV).
 C -> F (IN REF. 3).
 F -> T (IN REF. 3).
 MAJOR ALLERGEN I POLYPEPTIDE CHAIN 2.
 N-LINKED (GLCNAC. . .).
 TTISSKDCMGAEVQNTVEKLNTLGR -> PSTNIAMVK
 QERTP (in isoform 3).
 /FTID=VSP_004248.

Query Match 0.6%; Score 7; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163
 DB 4 ALLVLAL 10

RESULT 48

HDEA_ECOLI
 ID HDEA_ECOLI STANDARD; PRT; 110 AA.
 AC P26604;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein hdeA precursor (10K-S protein).
 GN HDEA OR B3510 OR Z4922 OR ECS4390 OR SF3544 OR S4223.
 OS Escherichia coli,
 OS Escherichia coli O157:H7, and

OS *Shigella flexneri*.
 CC Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562, 83334, 623;
 RN [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-40.
 RP SPECIES=E.coli; STRAIN=K12;
 RC MEDLINE=93204884; PubMed=8455549;
 RX Yoshida T., Ueguchi C., Yamada H., Mizuno T.;
 RA "Function of the Escherichia coli nucleoid protein, H-NS: molecular
 RT analysis of a subset of proteins whose expression is enhanced in a
 RT hns deletion mutant";
 RL Mol. Gen. Genet. 237:113-122(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=K12 / MG1655;
 RC MEDLINE=94316500; PubMed=8041620;
 RX Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RA "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes";
 RN Nucleic Acids Res. 22:2576-2586(1994).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RC MEDLINE=21074935; PubMed=11206551;
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21152231; PubMed=11258796;
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 RN [5]
 RN GENE NAME.
 RP SPECIES=E.coli;
 RC MEDLINE=94064579; PubMed=8244952;
 RX Yoshida T., Ueguchi C., Mizuno T.;
 RA "Physical map location of a set of Escherichia coli genes (hde) whose
 RT expression is affected by the nucleoid protein H-NS";
 RL J. Bacteriol. 175:7747-7748(1993).
 RN [6]
 RN SEQUENCE OF 22-41.
 RP SPECIES=E.coli; STRAIN=K12 / W3110;
 RC MEDLINE=94064579; PubMed=8244952;
 RX Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
 RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
 RA Hochstrasser D.F.;
 RL Submitted (SEP-1994) to Swiss-Prot.
 RN [7]
 RN SEQUENCE OF 22-33.
 RP SPECIES=E.coli; STRAIN=K12 / BMG2;
 RC MEDLINE=97443975; PubMed=9298466;
 RX Link A.J., Robison K., Church G.M.;
 RA "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12";
 RL Electrophoresis 18:1259-1313(1997).
 RN [8]
 RN SEQUENCE OF 22-31.
 RP SPECIES=E.coli; STRAIN=K12;
 RC MEDLINE=99085675; PubMed=9868784;
 RX Wasinger V.C., Humphrey-Smith I.;
 RA "Small genes/gene-products in Escherichia coli K-12";
 RT

FEMS Microbiol. Lett. 169:375-382(1998).
 RN [9]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 Yu J.;
 RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [10]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RC MEDLINE=2250274; PubMed=12704152;
 RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Pournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RA "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T";
 RL Infect. Immun. 71:2775-2786(2003).
 RN [11]
 RN MASS SPECTROMETRY.
 RP SPECIES=E.coli; STRAIN=ATCC 15597;
 RC MEDLINE=22067786; PubMed=12071744;
 RX Reid G.E., Shang H., Hogan J.M., Lee G.U., McLuckey S.A.;
 RA "Gas-phase concentration, purification, and identification of whole
 RT proteins from complex mixtures";
 RL J. Am. Chem. Soc. 124:7353-7362(2002).
 RN [12]
 RN X-RAY CRYSTALLOGRAPHY.
 RP SPECIES=E.coli;
 RC MEDLINE=98400493; PubMed=9731767;
 RX Yang F., Gustafson K.R., Boyd M.R., Wlodawer A.;
 RA "Crystal structure of Escherichia coli HdeA";
 RL Nat. Struct. Biol. 5:763-764(1998).
 RN [13]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RP SPECIES=E.coli;
 RC MEDLINE=20090957; PubMed=10623550;
 RX Gajiwala K.S., Burley S.K.;
 RA "HdeA, a periplasmic protein that supports acid resistance in
 RT pathogenic enteric bacteria";
 RL J. Mol. Biol. 295:605-612(2000).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -1- MASS SPECTROMETRY: MW=9742; METHOD=Electrospray; RANGE=22-110.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC -----
 DR ENBL; D11109; BAA01883.1; --
 DR ENBL; U00039; AAB18486.1; --
 DR ENBL; AE000427; AAC76535.1; --
 DR ENBL; AE005576; AAG58651.1; --
 DR ENBL; AF002565; BAB37813.1; --
 DR ENBL; AE015362; AAN44999.1; --
 DR ENBL; AE016992; AAP19187.1; --
 DR PIR; F91177; F91177.
 DR PIR; G86023; G86023.
 DR PIR; S30268; S30268.
 DR PDB; 1BG8; 10-SEP-98.
 DR PDB; 1DJ8; 10-DEC-99.
 DR SWISS-2DPAGE; P26604; COLI.
 DR EcoGene; EG11398; hdeA.
 DR Periplasmic; Signal; 3D-structure; Complete proteome.
 KW

```

FT SIGNAL      1 21
FT CHAIN      22 110
FT DISULFID   39 87
FT HELIX      34 36
FT STRAND     38 38
FT STRAND     39 43
FT TURN       44 44
FT TURN       47 49
FT HELIX      50 60
FT TURN       61 64
FT HELIX      66 68
FT HELIX      73 88
FT TURN       89 89
FT TURN       91 92
FT STRAND     94 94
FT HELIX      95 107
SQ SEQUENCE   110 AA; 11858 MW; 063262C4863FA2E9 CRC64;

Query Match      0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLL 1156
DB 9 LGGILLLL 15

RESULT 49
INSI_RAT
ID INS1_RAT STANDARD; PRT; 110 AA.
AC P01322;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin 1 precursor.
GN INSI OR INS-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RX SEQUENCE FROM N.A.
RP MEDLINE=80045034; PubMed=498283;
RA Cordell B., Bell G.I., Fischer E., Denoto F.M., Ullrich A.,
RA Pictet R.L., Rutter W.J., Goodman H.M.;
RA "Isolation and characterization of a cloned rat insulin gene.";
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,
RA Tizard R.;
RA Cell 18:533-543 (1979).
RN (2)
RX SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=80045035; PubMed=498284;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,
RA Tizard R.;
RA "The structure and evolution of the two nonallelic rat preproinsulin
genes.";
RN (3)
RX SEQUENCE FROM N.A.
RP MEDLINE=80240379; PubMed=6249167;
RA Lomedico P.T., Rosenthal N., Kolodner R., Efstratiadis A.,
RA Gilbert W.;
RA "The structure of rat preproinsulin genes.";
RN (4)
RX SEQUENCE OF 25-54 AND 90-110.
RP MEDLINE=70067613; PubMed=4311938;
RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margoliash E.,
RA Aten B., Oyer P.E.;
RA "Proinsulin and the biosynthesis of insulin.";
RN (5)
RX Recent Prog. Horm. Res. 25:207-282 (1969).
RP SEQUENCE OF 57-87.
RX STRAIN=New Zealand white; TISSUE=Pancreas;
RX MEDLINE=94179230; PubMed=8132571;

```

Tager H.S., Steiner D.F.;
 "Primary structures of the proinsulin connecting peptides of the rat
 and the horse.";

J. Biol. Chem. 247:7936-7940 (1972).
 [6]
 SEQUENCE OF 57-87, AND REVISIONS.
 MEDLINE=72177385; PubMed=4554104;
 Markussen J., Sundby F.;
 "Rat-proinsulin C-peptides. Amino-acid sequences.";

Eur. J. Biochem. 25:153-162 (1972).
 -!- FUNCTION: Insulin decreases blood glucose concentration. It
 increases cell permeability to monosaccharides, amino acids and
 fatty acids. It accelerates glycolysis, the pentose phosphate
 cycle, and glycogen synthesis in liver.
 -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 disulfide bonds.
 -!- SUBCELLULAR LOCATION: Secreted.
 -!- SIMILARITY: Belongs to the insulin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; V01242; CAA24559.1; -
 EMBL; J00747; AAA41442.1; -
 EMBL; M25584; AAA41439.1; -
 PIR; A90788; IPRT1.
 HSP; P01308; 1A7F.
 InterPro: IP0004825; Ins/IGF/relax.
 Pfam; PF00049; Insulin; 1.
 PRINTS; PR00277; INSULINB.
 SMART; SM00078; ILGF; 1.
 PROSITE; PS00262; INSULIN; 1.
 Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.

SIGNAL 1 24
 CHAIN 25 54 INSULIN 1 B CHAIN.
 PROPEP 57 87 INSULIN 1 C PEPTIDE.
 CHAIN 90 110 INSULIN 1 A CHAIN.
 DISULFID 31 96 INTERCHAIN.
 DISULFID 43 109 INTERCHAIN.
 DISULFID 95 100
 SEQUENCE 110 AA; 12420 MW; 51D606DA54AE3533 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
 DB 10 LLALLVL 16

RESULT 50
 INS_RABIT
 ID INS_RABIT STANDARD; PRT; 110 AA.
 AC P01311;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Insulin precursor.
 GN INS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX STRAIN=New Zealand white; TISSUE=Pancreas;
 RX MEDLINE=94179230; PubMed=8132571;

```

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
RA Menon R.K., Zahm D.S.;
RT "Insulin gene expression and insulin synthesis in mammalian neuronal
RT cells.";
RL J. Biol. Chem. 269:8445-8454(1994).
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949593;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RL Am. J. Med. 40:662-666(1966).
RN [3]
RP SEQUENCE OF 56-110 FROM N.A.
RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U03610; AAA19033.1; -.
DR EMBL; M61153; AAA17540.1; -.
DR PIR; A53438; INRB.
DR HSP; P01308; ILYM.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 83 E -> Y (IN REF. 3).
SQ SEQUENCE 110 AA; 11938 MW; 82D2975B85D77FA8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 51
IGF_MXGL
ID -IGF_MXGL STANDARD; PRT; 139 AA.
AC P22618;
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin-like growth factor precursor (IGF) (Fragment).
OS Myxine glutinosa (Atlantic hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxinae; Myxine.
NCBI_TaxID=7769;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91115860; PubMed=989990;
RA Nagamatsu S., Chan S.J., Falkner S., Steiner D.F.;
RT "Evolution of the insulin gene superfamily. Sequence of a
RT preproinsulin-like growth factor cDNA from the Atlantic hagfish.";
RL J. Biol. Chem. 266:2397-2402(1991).
CC -!- FUNCTION: The insulin-like growth factors, isolated from plasma,
CC are structurally and functionally related to insulin but have a
CC much higher growth-promoting activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M57735; AAA49265.1; -.
DR PIR; A39612; A38612.
DR HSP; P01344; IGF2.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Mitogen; Growth factor; Signal.
FT NON_TER 1 1
FT SIGNAL <1 38
FT CHAIN 39 139 INSULIN-LIKE GROWTH FACTOR.
FT DOMAIN 39 67 B.
FT DOMAIN 68 82 C.
FT DOMAIN 83 103 A.
FT DOMAIN 104 113 D.
FT DOMAIN 114 139 E.
SQ SEQUENCE 139 AA; 16087 MW; 2FC888C8D074FAC1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALL 1159
Db 28 LLLALL 34

RESULT 52
LAMP_PETWA
ID LAMP_PETWA STANDARD; PRT; 139 AA.
AC P33575; P33576;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lamprin 0.9 precursor (Cartilage matrix protein).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
NCBI_TaxID=7757;
RN [1]_TaxID=7757;
RP SEQUENCE FROM N.A. (ISOFORMS 0.9-10 AND 0.9-12), AND SEQUENCE OF
RP 20-44.
RC TISSUE=Cartilage;
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
CC -!- FUNCTION: Self-aggregating protein that makes part of the soluble

```

```

CC form of lamprin.
CC -!- SUBUNIT: The polymeric lamprin chains self-aggregate to form
CC fibers and have secondary structures particularly rich in beta-
CC sheets and in beta-turns.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=0.9-12;
CC IsoId=P33575-1; Sequence=Displayed;
CC Name=0.9-10;
CC IsoId=P33575-2; Sequence=VSP_004302;
CC Note=No experimental confirmation available;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05925; AAA49269.1; -;
CC EMBL; L05924; AAA49268.1; -;
CC PIR; C45051; C45051.
CC Repeat; Connective tissue; Signal; Alternative splicing.
CC FT SIGNAL 1 19 LAMPRIN 0.9.
CC FT CHAIN 20 139 8 X 5 AA APPROXIMATE REPEATS.
CC FT DOMAIN 42 110 1.
CC FT REPEAT 42 46 2.
CC FT REPEAT 47 51 3.
CC FT REPEAT 52 56 4.
CC FT REPEAT 57 61 5.
CC FT REPEAT 62 66 6.
CC FT REPEAT 67 71 7.
CC FT REPEAT 92 96 8.
CC FT REPEAT 106 110 Missing (in isoform 0.9-10).
CC FT VARSPLIC 86 104 /FTId=VSP_004302.
CC SQ SEQUENCE 139 AA; 13257 MW; E249ABB7A6CEC7C3 CRC64;
CC Query Match 0.6%; Score 7; DB 1; Length 139;
CC Best Local Similarity 100.0%; Pred. No. 77;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1157 ALLVIAL 1163
CC Db 7 ALLVIAL 13
CC
CC RESULT 53
CC HEAD CHICK STANDARD; PRT; 141 AA.
CC AC P02001;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Hemoglobin alpha-D chain.
CC GN HBA
CC OS Gallus gallus (Chicken).
CC OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
CC OC Gallus.
CC OX NCBI_TaxID=9031;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=White leghorn;
CC RX MEDLINE=92020223; PubMed=1656392;
CC RA Lewis W., Lee J.D., Dodgson J.B.;
CC RT "Adult chicken alpha-globin gene expression in transfected Q76 quail
CC cells: evidence for a negative regulatory element in the alpha D gene
CC region.";
CC RL Nucleic Acids Res. 19:5321-5329(1991).
CC RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=83161047; PubMed=6300093;
RA Dodgson J.B., Engel J.D.;
RT "The nucleotide sequence of the adult chicken alpha-globin genes.";
RL J. Biol. Chem. 258:4623-4629(1983).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82082384; PubMed=6273837;
RA Dodgson J.B., McCune K.C., Rusling D.J., Krust A., Engel J.D.;
RT "Adult chicken alpha-globin genes alpha A and alpha D: no anemic
RT shock alpha-globin exists in domestic chickens.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:5998-6002(1981).
RN [4]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=76189993; PubMed=1225908;
RA Takei H., Ota Y., Wu K.C., Kiyohara T., Matsuda G.;
RT "Amino acid sequence of the alpha chain of chicken AI hemoglobin.";
RL J. Biochem. 77:1345-1347(1975).
RN [5]
RP SEQUENCE OF 1-63 AND 94-122.
RC TISSUE=Embryo;
RX MEDLINE=82098109; PubMed=7054172;
RA Chapman B.S., Hood L.E., Tobin A.J.;
RT "Minor early embryonic chick hemoglobin M. Amino acid sequences of
RT the epsilon and alpha D chains.";
RL J. Biol. Chem. 257:651-658(1982).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=95156923; PubMed=10037733;
RA Knapp J.E., Oliveira M.A., Xie Q., Ernst S.R., Riggs A.F.,
RA Hackert M.B.;
RT "The structural and functional analysis of the hemoglobin D component
RT from chicken.";
RL J. Biol. Chem. 274:6411-6420(1999).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the
CC various peripheral tissues.
CC -!- SUBUNIT: Tetramer of two alpha-D chains and two beta chains.
CC The component D forms dimers of tetramers upon deoxygenation.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
CC hemoglobin component, called hemoglobin d, which is expressed in
CC late embryonic and adult life.
CC -!- MISCELLANEOUS: REF.5 CHAIN WAS ISOLATED FROM HBM, THE LEAST
CC ABUNDANT OF THE FOUR EARLY CHICK HEMOGLOBINS.
CC -!- SIMILARITY: Belongs to the globin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59989; CAA42605.1; -;
CC EMBL; V00411; CAA23702.1; -;
CC EMBL; M15378; AAA48584.1; ALT_SEQ.
CC EMBL; J00853; AAA48800.1; -;
CC PIR; B92421; HACH1.
CC PDB; 1HBR; 26-MAR-99.
CC InterPro; IPR002336; Alpha_haem.
CC InterPro; IPR000971; Globin.
CC Pfam; PF00042; globin; 1.
CC PRINTS; PR00612; ALPHAHAE.
CC PROSITE; PS01033; GLOBIN; 1.
CC Heme; Oxygen transport; transport; Erythrocyte;
CC 3D-structure.
CC FT METAL 58 58 IRON (HEME DISTAL LIGAND).
CC FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
CC FT CONFLICT 16 16 K -> R (IN REF. 1).
CC FT CONFLICT 107 107 V -> C (IN REF. 3 AND 5).
CC FT HELIX 4 17
CC FT TURN 18 19

```

PT HELIX 21 35
PT HELIX 37 42
FT TURN 44 51
FT TURN 50 51
FT HELIX 53 71
FT TURN 72 74
FT HELIX 76 79
FT HELIX 81 88
FT TURN 89 91
FT TURN 95 95
FT HELIX 96 113
FT HELIX 114 116
FT HELIX 119 136
FT TURN 137 137
SQ SEQUENCE 141 AA; 15695 MW; 1FE42696B7B5384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 625 GAGNAV 631
|||||
Db 64 GAGNAV 70

RESULT 54
HBAD MELGA STANDARD; PRT; 141 AA.
AC P81024;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-D chain.
GN HBAD.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
CX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE.
RX MEDLINE=96066279; PubMed=7576241;
RA Echichi Y., Ikehara T., Kayo S., Echichi T., Takei H.;
RT "Amino acid sequence of alpha- and beta-polypeptide chains of turkey
Meleagris gallopavo hemoglobin.";
RL Biol. Chem. Hoppe-Seyler 376:437-440 (1995).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the
various peripheral tissues.
CC -!- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
late embryonic and adult life.
CC -!- SIMILARITY: Belongs to the globin family.
DR HSP; P02001; IHR.
DR InterPro; IPR002338; Alpha haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PS00612; ALPHAHAE.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15665 MW; 59942696887954E7 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 625 GAGNAV 631
|||||
Db 64 GAGNAV 70

RESULT 55
MK_CHICK STANDARD; PRT; 142 AA.
AC P24052;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Midkine precursor (Retinoic acid-induced heparin-binding protein)
(RI-HB).
DE (RI-HB).
GN RIHB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 80-85; 112-118 AND 121-133.
RX TISSUE=Embryo;
RX MEDLINE=91207359; PubMed=2018506;
RA Urios P., Duprez D., le Caer J.-P., Courtois Y., Vigny M., Laurent M.;
RT "Molecular cloning of RI-HB, a heparin binding protein regulated by
retinoic acid.";
RT retinoic acid.";
RL Biochem. Biophys. Res. Commun. 175:617-624 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95010085; PubMed=7925417;
RA Duprez D., Treaguer J., Pecqueur C., Vigny M.R.;
RT "Organisation and promoter activity of the
retinoic-acid-induced-heparin-binding (RIHB) gene.";
RL Eur. J. Biochem. 224:931-941 (1994).
RN [3]
RP SEQUENCE OF 22-77.
RX MEDLINE=91128406; PubMed=1993066;
RA Raulais D., Lagente-Chevallier O., Guettet C., Duprez D., Courtois Y.,
Vigny M.;
RT "A new heparin binding protein regulated by retinoic acid from chick
embryo.";
RL Biochem. Biophys. Res. Commun. 174:708-715 (1991).
RN [4]
RP SEQUENCE OF 22-77.
RX STRAIN=White leghorn; TISSUE=Embryo;
RX MEDLINE=90108010; PubMed=2558016;
RA Vigny M., Raulais D., Puzenat N., Duprez D., Hartman M.P.,
Jeanny J.C., Courtois Y.;
RT "Identification of a new heparin-binding protein localized within
chick basement membranes.";
RL Eur. J. Biochem. 186:733-740 (1989).
CC -!- FUNCTION: Has mitogenic activity, and neurite extension activity
for PC12 cells.
CC -!- SUBCELLULAR LOCATION: BASEMENT MEMBRANES IN EARLY EMBRYONIC
TISSUES, AND CELL SURFACE OF NEUROECTODERMAL CELLS.
CC -!- DEVELOPMENTAL STAGE: Essentially expressed during embryogenesis.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: Belongs to the pleiotrophin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; M61754; -; NOT ANNOTATED_CDS.
DR EMBL; X76482; CAA54020.1; -;
DR PIR; JT0573; JT0573.
DR HSP; P21741; IMKC.
DR InterPro; IPR000762; PTN MK.
DR Pfam; PF01091; PTN MK C; 1.
DR Pfam; PF05196; PTN MK N; 1.
DR PRINTS; PR00269; PTN MIDKINE.
DR ProDom; PD005592; PTN MK; 1.

DR SMART: SMO0193; PTN: 1.
 DR PROSITE: PS00619; PTN_MK_1; 1.
 DR PROSITE: PS00620; PTN_MK_2; 1.
 KW Growth factor; Mitogen; Differentiation; Heparin-binding; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 142 MIDKINE.
 FT DISULFID 36 60 BY SIMILARITY.
 FT DISULFID 44 69 BY SIMILARITY.
 FT DISULFID 51 73 BY SIMILARITY.
 FT DISULFID 83 115 BY SIMILARITY.
 FT DISULFID 93 125 BY SIMILARITY.
 FT CONFLICT 89 89 S -> R (IN REF. 2).
 FT CONFLICT 93 93 C -> G (IN REF. 2).
 SQ SEQUENCE 142 AA; 15579 MW; 9D05CAF9B558451B CRC64;

Query Match 0.6%; Score 7; DB 1; Length 142;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
 Db 7 LLLALL 13

RESULT 56
 BIK_HUMAN
 ID BIK_HUMAN STANDARD; PRT; 160 AA.
 AC Q1323; Q16592;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bcl-2 interacting killer (Apoptosis inducer NBK) (BP4) (BIP1).
 GN BIK OR NBK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=B-cell;
 RX MEDLINE=96068922; PubMed=7478623;
 RA Boyd J.M., Gallo G.J., Elangovan B., Houghton A.B., Malstrom S.,
 RA Avery B.J., Ebb R.G., Subramanian T., Chittenden T., Lutz R.J.,
 RA Chinnadurai G.
 RT "Bik" a novel death-inducing protein shares a distinct sequence motif
 RT with Bcl-2 family proteins and interacts with viral and cellular
 RT survival-promoting proteins."
 RL Oncogene 11:1921-1928(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96413338; PubMed=8816500;
 RA Han J., Sabbatini P., White E.;
 RT "Induction of apoptosis by human NbK/Bik, a BH3-containing protein
 RT that interacts with E1B 19K."
 RL Mol. Cell. Biol. 16:5857-5864(1996).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Lymphoid;
 RA Pun K.-T., Farrow S.N., Raven T., Wride C.J., White J.H.M., Brown R.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99431905; PubMed=10500065;
 RA Castells A., Ino Y., Louis D.N., Ramesh V., Gusella J.F., Rustgi A.K.;
 RT "Mapping of a target region of allelic loss to a 0.5-cM interval on
 RT chromosome 22q13 in human colorectal cancer."
 RL Gastroenterology 117:831-837(1999).
 RN [5]
 RN SEQUENCE FROM N.A.
 RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Graffham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashrafi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA McElay J.N., McLaren S., McMuray A.A., Milne S.A., Mortimore B.J.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen P., Chu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
 RA Koef I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Sudar M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edlmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Keda D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22."
 RL Nature 402:489-495(1999).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Fahy J., Helton E., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.
 RX MEDLINE=96091131; PubMed=8521816;

```

RA Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J.,
RA Elangovan B., Chinnadurai G., Lutz R.J.;
RA "A conserved domain in Bak, distinct from BH1 and BH2, mediates cell
RT death and protein binding functions.";
RT EMBO J. 14:5589-5596(1995)
CC
CC -!- FUNCTION: Accelerates programmed cell death. Binding to the
CC apoptosis repressors Bcl-X(L), BHRF1, Bcl-2 or its adenovirus
CC homolog E1B 19k protein suppresses this death-promoting activity.
CC Does not interact with BAX.
CC
CC -!- SUBCELLULAR LOCATION: Around the nuclear envelope, and in
CC cytoplasmic membranes.
CC
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and
CC BAX for their pro-apoptotic activity and by their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U34584; AAC50413.1; -
CC DR EMBL; U49730; AAC79124.1; -
CC DR EMBL; X89986; CAA62013.1; -
CC DR EMBL; AF174424; AAF01156.1; -
CC DR EMBL; AF174421; AAF01156.1; JOINED.
CC DR EMBL; AF174422; AAF01156.1; JOINED.
CC DR EMBL; AF174423; AAF01156.1; JOINED.
CC DR EMBL; AV245248; AAF01089.1; -
CC DR EMBL; AL022237; CAA18260.2; -
CC DR EMBL; BC001599; AAR01599.1; -
CC DR PIR; S58214; S58214.
CC DR Genew; HGNC:1051; BIK.
CC DR MIM; 603392; -
CC DR GO; GO:0008632; P:apoptotic program; TAS.
CC DR GO; GO:0006917; P:induction of apoptosis; TAS.
CC DR InterPro; IPR000712; Bcl2_BH.
CC DR PROSITE; PS01259; BH3; 1.
CC KW Apoptosis; Transmembrane.
CC FT DOMAIN 57 71
CC FT TRANSMEM 136 156
CC FT DOMAIN 137 158
CC FT CONFLICT 149 150 PL -> LP (IN REF. 1).
CC FT SEQUENCE 160 AA; 18016 MW; 89034F4443F5A136 CRC64;
CC
CC Query Match 0.6%; Score 7; DB 1; Length 160;
CC Best Local Similarity 100.0%; Pred. No. 86;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
QY 1153 LLLALL 1159
DB 141 LLLALL 147
CC
CC RESULT 57
CC NUOE AQUAE
CC ID NUOE AQUAE STANDARD; PRT; 160 AA.
CC AC 066842;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE NADH-quinone oxidoreductase chain E (EC 1.6.99.5) (NADH dehydrogenase
CC I, chain E) (NDH-1, chain E).
CC GN NUOE OR AQ_574.
CC OS Aquifex aeolicus.
CC OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
CC OX NCBI_TaxID=63363;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=VF5;

```

```

RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shear M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RT Nature 392:353-358(1998).
CC
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (By similarity).
CC
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC
CC -!- COFACTOR: Binds 1 2Fe-2S cluster (Potential).
CC
CC -!- SIMILARITY: Belongs to the complex 1 24 kDa subunit family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000696; AAC06799.1; -
CC DR PIR; F70351; F70351.
CC DR InterPro; IPR002023; Cmplx1_24kDa.
CC DR Pfam; PF01257; complex_24kD; 1.
CC DR PRODOM; PDC03859; Cmplx1_24kDa; 1.
CC DR PROSITE; PS01099; COMPLEX_24k; 1.
CC KW Oxidoreductase; NAD; Quinone; Metal-binding; Iron-sulfur; Iron;
CC 2Fe-2S; Complete proteome.
CC FT METAL 86 86
CC FT METAL 91 91
CC FT METAL 127 127
CC FT METAL 131 131
CC FT SEQUENCE 160 AA; 18550 MW; 425D81995A491B35 CRC64;
CC
CC Query Match 0.6%; Score 7; DB 1; Length 160;
CC Best Local Similarity 100.0%; Pred. No. 86;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
QY 413 BPPEELK 419
DB 7 BPPEELK 13
CC
CC RESULT 58
CC TCH2 ARATH
CC ID TCH2 ARATH STANDARD; PRT; 161 AA.
CC AC P25070; O22592;
CC DT 01-MAY-1992 (Rel. 22, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Calmodulin-related protein 2, touch-induced.
CC GN TCH2 OR ATSG3770 OR K22F20.1.
CC OS Arabidopsis thaliana (Mouse-ear cress).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CC OX NCBI_TaxID=3702;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=cv. Columbia;
CC RA Johnson K.A., Braam J.;
CC RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC
CC [2]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=cv. Columbia;
CC RC MEDLINE=98087489; PubMed=9872454;
CC RX Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
CC Tabata S.;

```

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned p1 and TAC clones."
 RL DNA Res. 5:297-308(1998).
 RN [3]
 RP SEQUENCE OF 26-70 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=90150263; PubMed=2302732;
 RA Braam J., Davis R.W.;
 RT "Rain-, wind-, and touch-induced expression of calmodulin and
 RT calmodulin-related genes in Arabidopsis."
 RL Cell 60:357-364(1990).
 RN [4]
 RP 3D-STRUCTURE MODELING OF 7-158.
 RX MEDLINE=97189489; PubMed=9037719;
 RA Khan A.R., Johnson K.A., Braam J., James M.N.G.;
 RT "Comparative modeling of the three-dimensional structure of the
 RT calmodulin-related TCH2 protein from Arabidopsis."
 RL Proteins 27:144-153(1997).
 CC -!- INDUCTION: By rain-, wind-, and touch (thigmomorphogenesis).
 CC -!- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AF026473; BAB2713.1; -
 CC PIR; ABO16873; BAB10353.1; -
 CC PDB; C34569; C34669.
 CC DR LAVU; 28-JAN-98.
 CC InterPro; IPR003293; Calflagin.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF00036; efhand; 4.
 CC PRINTS; PR01362; CALFLAGIN.
 CC ProDom; PD000012; EF-hand; 2.
 CC SMART; SM00054; EFH; 4.
 CC PROSITE; PS00018; EF_HAND; 4.
 CC Calcium-binding; Repeat; 3D-structure.
 CC CA_BIND 26 37 EF-HAND 1 (POTENTIAL).
 CC CA_BIND 62 73 EF-HAND 2 (POTENTIAL).
 CC CA_BIND 103 114 EF-HAND 3 (POTENTIAL).
 CC CA_BIND 139 150 EF-HAND 4 (POTENTIAL).
 CC CONFLICT 54 54 T -> Y (IN REF. 3).
 CC HELIX 12 25
 CC TURN 27 28
 CC STRAND 33 33
 CC HELIX 35 44
 CC TURN 45 45
 CC HELIX 51 61
 CC TURN 63 64
 CC STRAND 69 69
 CC HELIX 71 102
 CC TURN 104 105
 CC STRAND 110 110
 CC HELIX 112 121
 CC TURN 122 123
 CC HELIX 128 138
 CC STRAND 146 146
 CC HELIX 148 156
 CC SEQUENCE 161 AA; 17546 MW; A7324A5C0B0CBDB9 CRC64;
 KW
 FT CA_BIND 26 37 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 62 73 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 103 114 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 139 150 EF-HAND 4 (POTENTIAL).
 FT CONFLICT 54 54 T -> Y (IN REF. 3).
 FT HELIX 12 25
 FT TURN 27 28
 FT STRAND 33 33
 FT HELIX 35 44
 FT TURN 45 45
 FT HELIX 51 61
 FT TURN 63 64
 FT STRAND 69 69
 FT HELIX 71 102
 FT TURN 104 105
 FT STRAND 110 110
 FT HELIX 112 121
 FT TURN 122 123
 FT HELIX 128 138
 FT STRAND 146 146
 FT HELIX 148 156
 SQ SEQUENCE 161 AA; 17546 MW; A7324A5C0B0CBDB9 CRC64;
 Query Match 0.6%; Score 7; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 487 VDIDGDG 493
 DB 138 VDIDGDG 144

RESULT 59
 YF87_METJA STANDARD; PRT; 171 AA.
 ID YF87_METJA
 AC Q58982;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M1587.
 GN M1587.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcales; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geobach N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.W., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii."
 RL Science 273:1058-1073(1996).
 CC -!- SIMILARITY: TO M.JANNASCHII MJ0417.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; U67599; AAB93616.1; -
 CC PIR; B64498; B64498.
 CC TIGR; MJ1587; -
 CC InterPro; IPR003141; PHP_N.
 CC Pfam; PF02231; PHP_N; 1.
 CC SMART; SM00481; POLIIFAC; 1.
 CC Hypothetical protein; Complete proteome.
 CC SEQUENCE 171 AA; 19198 MW; 7D4F7B936C5C7472 CRC64;
 KW
 FT PIR; B64498; B64498.
 FT TIGR; MJ1587; -
 FT InterPro; IPR003141; PHP_N.
 FT Pfam; PF02231; PHP_N; 1.
 FT SMART; SM00481; POLIIFAC; 1.
 FT Hypothetical protein; Complete proteome.
 FT SEQUENCE 171 AA; 19198 MW; 7D4F7B936C5C7472 CRC64;
 Query Match 0.6%; Score 7; DB 1; Length 171;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 344 ALGDRIF 350
 DB 123 ALGDRIF 129

RESULT 60
 YD2F_SCHPO STANDARD; PRT; 176 AA.
 ID YD2F_SCHPO
 AC Q10263;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Very hypothetical protein C56F8.15 in chromosome I.
 GN SPAC56F8.15.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brannner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Graven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RA "The complete genome sequence of *Chlorobium tepidum* TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).
CC -/- CATALYTIC ACTIVITY: D-erythro-1-(imidazol-4-yl)glycerol 3-
CC phosphate = 3-(imidazol-4-yl)-2-oxopropyl phosphate + H(2)O.
CC -/- PATHWAY: Histidine biosynthesis; sixth step.
CC -/- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -/- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL curation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
CC or send an email to license@isb-sib.ch).
CC

DR PFam: PF000475; IGBP: 1;
DR PRODom: PD002282; IGBP: 1;
DR PROSITE: PS00954; IGP_DEHYDRATASE_1; 1;
DR PROSITE; PS00955; IGP_DEHYDRATASE_2; 1;
DR Histidine biosynthesis; Lyase; Complete proteome.
KW Histicidine biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 194 AA; 21189 MW; D8725AAC9B83032 CRC64;

```

Query Match          0.6%; Score 7; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      341  IVDALGD 347
D5      80  IVDALGD 86
          |||||
RESULT 62

```

ID	STANDARD:	PRT:	201 AA.
AC	-Z2399 MOUSE		
AC	P24359;		
DT	01-MAR-1992 (Rel. 21, Created)		
DT	01-MAR-1992 (Rel. 21, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Zinc finger protein 239 (ZFP-239) (Zinc finger protein MOK-2).		
GN	ZNF239 OR ZFP239 OR MOK2 OR MOK-2.		

OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
FN	[1]
RP	SEQUENCE FROM N.A. PubMed=2104562;
RX	MEDLINE=90097859; Hamer D.;
RA	Ernault-Lange M., Kress M.,
RT	"A gene that encodes a protein consisting solely of zinc finger
RT	domains is preferentially expressed in transformed mouse cells.";
RL	Mol. Cell. Biol. 10:418-421(1990).
CC	-!- FUNCTION: May function as a transcription factor.
CC	-!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	-!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN TRANSFORMED
CC	MOUSE CELLS.
CC	-!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC	FINGER PROTEINS.
CC	-!- SIMILARITY: Contains 7 C2H2-type zinc fingers.
CC	-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M32057; AAA39729.1; --
 DR PIR: I57505; I57505.
 DR HSSP: P08047; I5P2.
 DR TRANSFAC: T00510; --
 DR MGD: MGI:1306812; Zfp239.
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam: PF00096; zf-C2H2; 7.
 DR PRINTS: PR00048; ZINC_FINGER.
 DR ProDom: PD000003; Znf_C2H2; 6.
 DR SMART: SM00355; Znf_C2H2; 7.
 DR PROSITE: PS00268; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 7.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 FT ZN_FING 6 28 C2H2-TYPE 1.
 FT ZN_FING 34 56 C2H2-TYPE 2.
 FT ZN_FING 62 84 C2H2-TYPE 3.
 FT ZN_FING 90 112 C2H2-TYPE 4.
 FT ZN_FING 118 140 C2H2-TYPE 5.
 FT ZN_FING 146 168 C2H2-TYPE 6.
 FT ZN_FING 174 196 C2H2-TYPE 7.
 SQ SEQUENCE 201 AA; 22832 MW; 57A18COAADAB1384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 959 FTRSSSL 965
 Db 15 FTRSSSL 21
 |||||
 |||||

RESULT 63
 GPBB_MOUSE
 ID GPBB_MOUSE STANDARD; PRT; 206 AA.
 AC P56400;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Platelet glycoprotein Ib beta chain precursor (GP-Ib beta) (GPtBB)
 DE (GPtB-beta).
 GN GPtBB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97403789; PubMed=9259114;
 RA Kitaguchi T., Murata M., Anbo H., Moriki T., Ikeda Y.;
 RT "Characterization of the gene encoding mouse platelet glycoprotein Ib
 beta.";
 RL Thromb. Res. 87:235-244(1997).
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,
 CC participates in the formation of platelet plugs by binding to von
 CC Willebrand factor, which is already bound to the subendothelium
 CC (By similarity).
 CC -!- SUBUNIT: GP-Ib alpha and beta are disulfide linked. GP-IX is
 CC complexed with the GP-Ib heterodimer via a non covalent linkage
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption
 CC of the macromolecular complex of GP-Ib with the platelet
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-

CC binding protein.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeat.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB001419; BAA22424.1; --
 DR MGD: MGI:107852; Gp1bb.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR00372; LRR_Nterm.
 DR Pfam: PF00560; LRR; 1.
 DR Pfam: PF01463; LRRCT; 1.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00013; LRRNT; 1.
 KW Signal; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;
 KW Platelet; Phosphorylation; Cell adhesion; Leucine-rich repeat.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 206 PLATELET GLYCOPROTEIN IB BETA CHAIN.
 FT DOMAIN 27 147 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 148 172 POTENTIAL.
 FT DOMAIN 173 206 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 60 83 LRR.
 FT CARBOHYD 66 66 N-LINKED (GLCNAC...) (POTENTIAL).
 FT MOD RES 191 191 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
 SQ SEQUENCE 206 AA; 21762 MW; AC4BCB4DFA226FID CRC64;

Query Match 0.6%; Score 7; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALL 1159
 Db 13 LLLALL 19
 |||||
 |||||

RESULT 64
 PSMB_SULSO
 ID PSMB_SULSO STANDARD; PRT; 208 AA.
 AC Q9UXF3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Proteasome beta subunit precursor (EC 3.4.25.1) (Multicatalytic
 DE endopeptidase complex beta subunit).
 GN PSMB OR SSO0766 OR C40_002.
 GN PSMB OR SSO0766 OR C40_002.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis B., Duguet M., Erasmo G., Faguy D.,
 RA Gaucherland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kushwaha N., Laflaur B., Medina N., Peng X., Penny S.L., She Q.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A., Sengen C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;

RA She Q., Singh R.K., Confalonieri P., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RA The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: The proteasome is a multicatalytic complex
CC which is characterized by its ability to cleave peptides with Arg,
CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
CC slightly basic pH (By similarity).
CC -!- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
CC specificity.
CC -!- SUBUNIT: Composed of two subunits, alpha and beta. The complex is
CC formed of four rings. The two outer rings are each composed of
CC seven alpha subunits. The two inner rings are each composed of
CC seven beta subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to peptidase family T1B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y18930; CAB57537.1; ALT_INIT.
CC EMBL: AE008700; AAK41063.1; ALT_INIT.
CC HSP: P28061; 1PXA.
CC MEROPS: T01.002; -.
CC InterPro: IPR000243; Pept T1A sub.
CC InterPro: IPR001353; Peptidase_T1.
CC Pfam: PF00227; Proteasome; 1.
CC PRINTS: PR00141; PROTEASOME.
CC PROSITE: PS00854; PROTEASOME_B; 1.
CC Proteasome; Hydrolase; Protease; Complete proteome;
KW Threonine protease.
FT PROPEP 1 14 APPARENTLY REMOVED IN MATURE FORM (BY
FT SIMILARITY).
FT CHAIN 15 208 PROTEASOME BETA SUBUNIT.
FT ACT_SITE 15 15 PROBABLE.
FT SEQUENCE 208 AA; 22937 MW; C940A8273AB25CB9 CRC64;
Query Match 0.6%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 686 TTTVGIR 692
Db 15 TTTVGIR 21
RESULT 65
KPTA PYRAE STANDARD; PRT; 213 AA.
AC Q823P2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable RNA 2'-phosphotransferase (EC 2.7.-.-).
GN KPTA OR PAR3647.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=IM2 / ATCC 51768 / DSM 7523;
RC MEDLINE=21664397; PubMed=11792869;
RX

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RA "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*
RA *aerophilum*.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- FUNCTION: Removes the 2'-phosphate from RNA via an intermediate in
CC which the phosphate is ADP-ribosylated by NAD followed by a
CC presumed transesterification to release the RNA and generate ADP-
CC ribose 1''-2''-cyclic phosphate (APPR>P). May function as an ADP-
CC ribosylase (By similarity).
CC -!- SIMILARITY: Belongs to the kpta / Tpt1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE009943; AAL55071.1; -.
CC HAMAP: MF_00299; 1.
CC InterPro: IPR002745; Pirans_Kpta/Tpt1.
CC Pfam: PF01885; PTS_2-RNA; 1.
CC ProDom: PD009619; P04transfrase; 1.
CC Transfrase; NAD; Complete proteome.
CC SEQUENCE 213 AA; 23452 MW; 12DF2AD12CB69636 CRC64;
Query Match 0.6%; Score 7; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 DVYKCPV 78
Db 3 DVYKCPV 9
RESULT 66
CH13 HUMAN STANDARD; PRT; 214 AA.
ID CH13 HUMAN
AC Q96KS9; Q9N3M3; Q9NSR0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C8orf13.
GN C8ORF13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RT "Transcript map of the KWE critical region on chromosome 8p22-p23.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=amygdala;
RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 83-214 FROM N.A.
RC TISSUE=amygdala;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SEC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

MBL; AF039265; AAC33542.1; --.

Query Match 0.6%; Score 7; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

or send an email to license@isb-sib.ch).

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC
CC
DR EMBL; AB043007; BAB19277.1; -
DR EMBL; AF277316; AAK69113.1; -
DR EMBL; BC006248; AAH06248.1; -
DR PIR; JC7587; JC7587.
DR Genew; HGNC:10676; SDF2L1.
DR MIM; 607551; -
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR003608; MIR.
DR Pfam; PF02815; MIR_3.
DR SMART; SM00472; MIR_3.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00919; MIR_3.
DR Endoplasmic reticulum; Signal; Repeat.
KW POTENTIAL.
FT SIGNAL 1 28
FT CHAIN 29 221
FT STROMAL CELL-DERIVED FACTOR 2-LIKE
FT PROTEIN 1.
FT DOMAIN 33 87
FT MIR 1.
FT DOMAIN 95 150
FT MIR 2.
FT DOMAIN 151 205
FT MIR 3.
FT SITE 218 221
FT PREVENT SECRETION FROM ER (POTENTIAL).
FT C -> R (IN REF. 2 AND 3).
FT L -> F (IN REF. 2 AND 3).
FT CONFLICT 94 94
FT CONFLICT 162 162
FT SEQUENCE 221 AA; 23511 MW; 23EBE14BE932EF CRC64;
SQ

Query Match 0.6%; Score 7; DB 1; Length 221;
Best Local Similarity 100.0%; Pred No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLALLV 1160
| | | | |
DB 17 LLALLV 23

RESULT 70
EMBP MOUSE ID_EMBP_MOUSE STANDARD; PRT; 223 AA.
AC Q61878;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP) (Proteoglycan
2, bone marrow).
GN PRG2 OR MBP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CBA/J; TISSUE=Liver;
RX MEDLINE=95403975; PubMed=7673718;
RA Larson K.A., Horton M.A., Madden B.J., Gleich G.J., Lee N.A.,
RA Lee J.J.;
RT "The identification and cloning of a murine major basic protein gene
expressed in eosinophils.";
RL J. Immunol. 155:3002-3012(1995).
CC -!- FUNCTION: Cytotoxin and helminthotoxin. MBP also induces
noncytolytic histamine release from basophils. It is involved in
antiparasitic defense mechanisms and immune hypersensitivity
reactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

MATCHES 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC
CC
DR EMBL; AB043007; BAB19277.1; -
DR EMBL; AF277316; AAK69113.1; -
DR EMBL; BC006248; AAH06248.1; -
DR PIR; JC7587; JC7587.
DR Genew; HGNC:10676; SDF2L1.
DR MIM; 607551; -
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR003608; MIR.
DR Pfam; PF02815; MIR_3.
DR SMART; SM00472; MIR_3.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00919; MIR_3.
DR Endoplasmic reticulum; Signal; Repeat.
KW POTENTIAL.
FT SIGNAL 1 28
FT CHAIN 29 221
FT STROMAL CELL-DERIVED FACTOR 2-LIKE
FT PROTEIN 1.
FT DOMAIN 33 87
FT MIR 1.
FT DOMAIN 95 150
FT MIR 2.
FT DOMAIN 151 205
FT MIR 3.
FT SITE 218 221
FT PREVENT SECRETION FROM ER (POTENTIAL).
FT C -> R (IN REF. 2 AND 3).
FT L -> F (IN REF. 2 AND 3).
FT CONFLICT 94 94
FT CONFLICT 162 162
FT SEQUENCE 221 AA; 23511 MW; 23EBE14BE932EF CRC64;
SQ

Query Match 0.6%; Score 7; DB 1; Length 221;
Best Local Similarity 100.0%; Pred No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLALLV 1160
| | | | |
DB 17 LLALLV 23

RESULT 70
EMBP MOUSE ID_EMBP_MOUSE STANDARD; PRT; 223 AA.
AC Q61878;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP) (Proteoglycan
2, bone marrow).
GN PRG2 OR MBP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CBA/J; TISSUE=Liver;
RX MEDLINE=95403975; PubMed=7673718;
RA Larson K.A., Horton M.A., Madden B.J., Gleich G.J., Lee N.A.,
RA Lee J.J.;
RT "The identification and cloning of a murine major basic protein gene
expressed in eosinophils.";
RL J. Immunol. 155:3002-3012(1995).
CC -!- FUNCTION: Cytotoxin and helminthotoxin. MBP also induces
noncytolytic histamine release from basophils. It is involved in
antiparasitic defense mechanisms and immune hypersensitivity
reactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

RESULTS 69
SDFL_HUMAN STANDARD; PRT; 221 AA.
AC Q9HCN8; Q9BR15;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Stromal cell-derived factor 2-like protein 1 precursor (SDF2 like
protein 1) (WP1-interacting protein 8).
DE SDF2L1.
GN Homo sapiens (Human).
OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
PP SEQUENCE FROM N.A.
RP TissueTestis;
RC MEDLINE=21092619; PubMed=11162531;
RR Fukuda S., Sumii M., Masuda Y., Takahashi M., Koike N., Teishima J.,
RRA Yasumoto H., Itamoto T., Asahara T., Dohi K., Kamiya K.;
RRA "Murine and human SDF2L1 is an endoplasmic reticulum stress-inducible
gene and encodes a new member of the Pnt/rf protein family.";
RRT Biochem. Biophys. Res. Commun. 280:407-414(2001).
RRL [2]
RN [2]
RP SEQUENCE FROM N.A.
RR TissueUterus;
RC Tissue=Uterus;
RR Honore B.;
RRT hpwpl-interacting protein 8.";
RRL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RR TissueLung;
RC MEDLINE=22388257; PubMed=12477932;
RR Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RRA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RALtschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RADiatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RABrownstein M.J., Udwin T.B., Toskuyuki S., Carninci P., Frange C.,
RARaha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RASak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RARichards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RAFahy J., Hellon E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.J., Green E.D., Dickson M.C.,
RALakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RAButterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RASchnerzh A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with high expression in
testis, moderate expression in the pancreas, spleen, prostate,
small intestine and colon. Very low expression is seen in brain
and skeletal muscle.
CC -!- SIMILARITY: Contains 3 MIR domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

or send an email to license@isb-sib.ch).

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC CC
DR EMBL; AB043007; BAB19277.1; -
DR EMBL; AF277316; AAK69113.1; -
DR EMBL; BC006248; AAH06248.1; -
DR PIR; JC7587; JC7587.
DR Genew; HGNC:10676; SDF2L1.
DR MIM; 607551; -
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR003608; MIR.
DR Pfam; PF02815; MIR_3.
DR SMART; SM00472; MIR_3.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00919; MIR_3.
KW Endoplasmic reticulum; Signal; Repeat.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 221 STROMAL CELL-DERIVED FACTOR 2-LIKE
FT PROTEIN 1.
FT DOMAIN 33 87 MIR 1.
FT DOMAIN 95 150 MIR 2.
FT DOMAIN 151 205 MIR 3.
FT SITE 218 221 PREVENT SECRETION FROM ER (POTENTIAL).
FT CONFLICT 94 94 C -> R (IN REF. 2 AND 3).
FT CONFLICT 162 162 L -> F (IN REF. 2 AND 3).
SQ SEQUENCE 221 AA; 23511 MW; 23EBE14BE9B92EF CRC64;

Query Match 0.6%; Score 7; DB 1; Length 221;
Best Local Similarity 100.0%; Pred No. 1.1e+02; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 1154 LLALLV 1160
DB 17 LLALLV 23

RESULT 70
EMBP MOUSE ID_EMBP_MOUSE STANDARD; PRT; 223 AA.
AC Q61878;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP) (Proteoglycan
2, bone marrow).
GN PRG2 OR MBP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CBA/J; TISSUE=liver;
RX MEDLINE=95403975; PubMed=7673718;
RA Larson K.A., Horton M.A., Madden B.J., Gleich G.J., Lee N.A.,
RA Lee J.J.;
RT "The identification and cloning of a murine major basic protein gene
expressed in eosinophils.";
RL J. Immunol. 155:3002-3012(1995).
CC -!- FUNCTION: Cycotoxin and helminthotoxin. MBP also induces
noncytolytic histamine release from basophils. It is involved in
antiparasitic defense mechanisms and immune hypersensitivity
reactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

MATCHES 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC CC
DR 1153 LLALLV 1159
DR 6 LLALLV 12

RESULT 69
SDFL_HUMAN SDFL_HUMAN STANDARD; PRT; 221 AA.
AC Q9HCN8; Q9BR15;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Stromal cell-derived factor 2-like protein 1 precursor (SDF2 like
protein 1) (WPPI-interacting protein 8).
DE SDF2L1.
GN Homo sapiens (Human).
OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RS TissueTestis;
RM MEDLINE=21092619; PubMed=11162531;
RR Fukuda S., Sumii M., Masuda Y., Takahashi M., Koike N., Teishima J.,
RRA Yasumoto H., Itamoto T., Asahara T., Dohi K., Kamiya K.;
RRA "Murine and human SDF2L1 is an endoplasmic reticulum stress-inducible
gene and encodes a new member of the Pnt/rf protein family.";
RRT Biochem. Biophys. Res. Commun. 280:407-414(2001).
RRL [2]
RRT [2]
RP SEQUENCE FROM N.A.
RS TissueUterus;
RM Honore B.;
RRA HPWPI-interacting protein 8.";
RRL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RS Tissue=Lung;
RM MEDLINE=22388257; PubMed=12477932;
RR Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RRA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RRA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RRA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RRA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RRA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RRA Brownstein M.J., Udwin T.B., Toskuyuki S., Carninci P., Frange C.,
RRA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RRA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RRA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RRA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RRA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RRA Whiting M., Madan A., Young A.J., Green E.D., Dickson M.C.,
RRA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RRA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RRA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with high expression in
testis, moderate expression in the pancreas, spleen, prostate,
small intestine and colon. Very low expression is seen in brain
and skeletal muscle.
CC -!- SIMILARITY: Contains 3 MIR domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

CC EMBL; L46768; AAA83027.1; -
DR HSP; P13727; IHRU.
DR MGD; MGI:103294; Prg2
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00770; EMAJORBASICP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 106
FT CHAIN 107 223
FT DOMAIN 124 223
FT DISULFID 126 221
FT DISULFID 198 213
SQ SEQUENCE 223 AA; 24255 MW; 7D6D946DCEADA00 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 71
EMBP_RAT
ID EMBP_RAT STANDARD; PRT; 227 AA.
AC Q63189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP).
GN PRG2 OR MBP
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
MEDLINE=56138543; PubMed=8547309;
RA Nitton T., Watanabe M., Okayama H., Misawa S., Isobe Y., Hayashi H.,
RA Mue S., Ohuchi K.;
RT "Cloning of cDNA for rat eosinophil major basic protein.";
RL Biochim. Biophys. Acta 1364:261-264(1995).
CC -!- FUNCTION: Cytochrome and helminthotoxin. MBP also induces
CC noncytolytic histamine release from basophils. It is involved in
CC antiparasitic defense mechanisms and immune hypersensitivity
CC reactions (by similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50568; BAA09129.1; -
DR PIR; S68150; S68150.
DR HSP; P13727; IHRU.
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00770; EMAJORBASICP.

```

```

DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 110
FT CHAIN 111 227
FT DOMAIN 128 227
FT DISULFID 130 225
FT DISULFID 202 217
FT DOMAIN 59 68
SQ SEQUENCE 227 AA; 25129 MW; 34F2B5514090DE82 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 72
EMBI_CAVPO
ID EMBI_CAVPO STANDARD; PRT; 233 AA.
AC P22032;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein 1 precursor (MBP-1).
GN MBP1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 115-162.
RC TISSUE=Eosinophil;
RX MEDLINE=91160746; PubMed=1705901;
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA Saito T., Tasaka K.;
RT "Sequencing and cloning of the cDNA of guinea pig eosinophil major
RT basic protein.";
RL FEBS Lett. 279:330-334 (1991).
RN [2]
PARTIAL SEQUENCE.
RX MEDLINE=94092714; PubMed=8268206;
RA Hashimoto Y., Nagaoka I., Yamashita T.;
RT "Purification of the antibacterial fragments of guinea-pig major
RT basic protein.";
RL Biochim. Biophys. Acta 1203:236-242 (1993).
CC -!- FUNCTION: MBP may play some important roles in the allergic
CC reactions and inflammations, since MBP is capable of releasing
CC histamine from mast cells and damaging the epithelial cells of
CC bronchial tubes. Antiparasitic and antibiotic.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90251; BAA14291.1; -
DR PIR; S13625; S13625.
DR HSP; P13727; IHRU.
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR

```

DR PRINTS; PRO0770; EMAJORBASICP.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin;
KW Multigene family.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 114 ACIDIC.
FT CHAIN 115 233 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 1.
FT DOMAIN 132 233 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 134 231 BY SIMILARITY.
FT DISULFID 208 223 BY SIMILARITY.
SQ SEQUENCE 233 AA; 26268 MW; C8D5E96D927C55C8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
DB 4 LLLALL 10

RESULT 73
YBEU ECOLI
ID YBEU ECOLI STANDARD; PRT; 235 AA.
AC P77427;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybeu.
GN YBEU OR B0648.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28377; AAA69134.1; -.
CC EMBL; AB000379; AAC76003.1; -.
CC PIR; E65082; E65082.
CC EcoGene; EG12987; YGSA.
CC InterPro; IPR007563; DUF554.
CC Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC or send an email to license@isb-sib.ch.
CC EMBL; AE000169; AAC73749.1; -.
CC EMBL; U82598; AAB40849.1; -.
DR EMBL; D90704; BAA35295.1; -.
DR EMBL; D90705; BAA35300.1; -.
DR PIR; F64799; F64799.
DR EcoGene; EG13651; YbeU.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 235 AA; 27022 MW; 9A1CAF5D6A0ECF1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 936 ELAAGSD 932
DB 202 ELAAGSD 208

RESULT 74
YOGA ECOLI
ID YOGA ECOLI STANDARD; PRT; 235 AA.
AC Q46831.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yga.
GN YOGA OR B2966.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28377; AAA69134.1; -.
CC EMBL; AB000379; AAC76003.1; -.
CC PIR; E65082; E65082.
CC EcoGene; EG12987; YGSA.
CC InterPro; IPR007563; DUF554.
CC Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1151 GGLLLA 1157
RR2_MAIZE
DB 194 GGLLLA 200

RESULT 75
RR2_MAIZE
ID _RR2_MAIZE STANDARD; PRT; 236 AA.
AC P16037;
DT 01-APR-1990 (Rel. 14, Created)
DI 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Zea mays (Maize).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
NC 1]
RN SEQUENCE FROM N.A.
RP MEDLINE=90175001; PubMed=2308853;
EX Igloi G.L., Meinke A., Doery I., Koessel H.;
RA "Nucleotide and derived amino acid sequence of rps2 from maize
RT chloroplasts.";
RL Nucleic Acids Res. 18:663-663(1990).
NC 2]
RN SEQUENCE FROM N.A.
RP STRAIN=CV; FR9CMSSR37; TISSUE=Leaf;
RX MEDLINE=90272437; PubMed=2140888;
RA Stahl D., Rodermeil S., Subramanian A.R., Bogorad L.;
RT "Nucleotide sequence of a 3.46 kb region of maize chloroplast DNA
RT containing the gene cluster rpoC2-rps2-atpI-atpH.";
RL Nucleic Acids Res. 18:3073-3074(1990).
NC 3]
RN SEQUENCE FROM N.A.
RP MEDLINE=95395841; PubMed=7666415;
EX Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RA "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
RL J. Mol. Biol. 251:614-628(1995).
NC 4]
RN SEQUENCE OF 1-36 FROM N.A.
RP MEDLINE=90340289; PubMed=2381419;
RA Igloi G.L., Meinke A., Doery I., Koessel H.;
RT "Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:
RT comparison between the derived protein primary structures from
RT various organisms with respect to functional domains.";
RL Mol. Gen. Genet. 221:379-394(1990).
CC -1- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17318; CAA35198.1; -
DR EMBL; X52270; CAA36512.1; -
DR EMBL; X86563; CAA60273.1; -
DR PIR; S08249; R3ZM2.
DR Gramene; P16037; -
DR MaizEDB; 66014; -
DR HAMAP; MF_00291; -; 1.
DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMs; TIGR01011; rpsB_bact; 1.

```

```

DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 236 AA; 26881 MW; 80A3F99A8955SFAD5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 236;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLD 170
DB 163 DIVIVLD 169

Search completed: September 21, 2004, 13:32:10
Job time : 34 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:22:25 ; Search time 74 Seconds
(without alignments)
5065.346 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPRLGVVAMSLWPQFT.....FSARRRRPGLDTPKVL 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp Vertebrate.*
- 14: sp Unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	662	55.7	823	4	Q8WY18
2	122	10.3	1188	11	Q7TQC3
3	45	3.8	823	11	Q8C884
4	15	1.3	1171	13	Q42094
5	11	0.9	288	11	Q8EM12
6	10	0.8	116	13	Q91670
7	10	0.8	399	16	Q8XSH5
8	9	0.8	164	4	Q8N112
9	9	0.8	315	16	Q8YCE3
10	9	0.8	316	16	Q8FVL4
11	9	0.8	400	6	Q8HY16
12	9	0.8	413	16	Q89LS8
13	9	0.8	497	13	Q73804
14	9	0.8	607	10	Q39775
15	9	0.8	780	13	Q06271
16	8	0.7	54	6	Q7S65

90 Q8p60 xanthomonas
 91 Q9p3u1 schizosacch
 92 Q8y57 brucella me
 93 Q8fy8 brucella su
 94 Q04611 arabidopsis
 95 Q8ewn6 mycoplasma
 96 Q9ey90 amycolators
 97 Q82120 streptomyce
 98 Q52642 pseudomonas
 99 Q51991 pseudomonas
 100 Q64584 rattus norv

ALIGNMENTS

RESULT 1
 Q8WY18
 ID Q8WY18 PRELIMINARY; PRT; 823 AA.
 AC Q8WY18.
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE MSTP018.
 GN MST018.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aorta;
 RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,
 RA Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,
 RA Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.F.,
 RA Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF111799; AAL39001.1;
 DR GO; GO:0008305; C:integrin complex; IEA.
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR Pfam; PF01839; FG-GAP; 3.
 DR PRINTS; PR01185; INTEGRINA.
 DR SMART; SM00191; Int alpha; 4.
 SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;
 Query Match 55.7%; Score 662; DB 4; Length 823;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 366 MSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKPPPELKNHGAYL 425
 DB 1 MSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKPPPELKNHGAYL 60
 QY 426 GYTVTSVSSRQRRVYVAGAPRNFHTKGVILFTMHNRLSLTHQAMRGQIGSYFGSBIT 485
 DB 61 GYTVTSVSSRQRRVYVAGAPRNFHTKGVILFTMHNRLSLTHQAMRGQIGSYFGSBIT 120
 QY 486 SVDIDGDGVTDLVLCAPMYNEGERGVYVELRQNFVYNGTLKDSHSYQNAFQSS 545
 DB 121 SVDIDGDGVTDLVLCAPMYNEGERGVYVELRQNFVYNGTLKDSHSYQNAFQSS 180
 QY 546 IASVRDLNDSYNDVVVGAPLEDNAGAIYIFHGFSGSLTKPKORITASELATGLQYFG 605
 DB 181 IASVRDLNDSYNDVVVGAPLEDNAGAIYIFHGFSGSLTKPKORITASELATGLQYFG 240
 QY 605 CSTHGOLDNEGLDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCKRSGRD 665
 DB 241 CSTHGOLDNEGLDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCKRSGRD 300
 QY 666 ATCLAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGDRFTNRAVLSSG 725

301 ATCLAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGDRFTNRAVLSSG 360
 QY 726 QELCERINPHVLDTDADYKVPVTFVSVEYSLEDPDHPGMLDDGWPPTLTVSVFPWNGCNEDE 785
 DB 361 QELCERINPHVLDTDADYKVPVTFVSVEYSLEDPDHPGMLDDGWPPTLTVSVFPWNGCNEDE 420
 QY 786 HCVPLDVLDAASDLPTAMEYQORVLRKPAQCSAYTISFDTVFIESTRORVAEATLE 845
 DB 421 HCVPLDVLDAASDLPTAMEYQORVLRKPAQCSAYTISFDTVFIESTRORVAEATLE 480
 QY 846 NRGENAYSTVLNISQSANLQFASLIQKEDSDSGSECVNEERLQKQVCNVSYPFFRAKAK 905
 DB 481 NRGENAYSTVLNISQSANLQFASLIQKEDSDSGSECVNEERLQKQVCNVSYPFFRAKAK 540
 QY 906 VAFRLDSEFSKSI FLHLEIELAAGSDSNERDSTKEDNAPLRPHLYKYEADVLETRSSSL 965
 DB 541 VAFRLDSEFSKSI FLHLEIELAAGSDSNERDSTKEDNAPLRPHLYKYEADVLETRSSSL 600
 QY 966 SHYEVKLNSSLERYDGIGPPESCIFRIQNLGLFPIHGMMKTIPTIATRSNRLKLKRD 1025
 DB 601 SHYEVKLNSSLERYDGIGPPESCIFRIQNLGLFPIHGMMKTIPTIATRSNRLKLKRD 660
 QY 1026 LTDEANTSCNIWGNSTERYPTVEEDLRRAPQLNHSNDVVSINCNILVYVNOEINPHLL 1085
 DB 661 LTDEANTSCNIWGNSTERYPTVEEDLRRAPQLNHSNDVVSINCNILVYVNOEINPHLL 720
 QY 1086 GNLWLSLKALKYKSMKIMVNAALQROFHSPFIFREEDPSRQI 1128
 DB 721 GNLWLSLKALKYKSMKIMVNAALQROFHSPFIFREEDPSRQI 763
 RESULT 2
 Q7TQC3
 ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.
 AC Q7TQC3.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE All integrin.
 GN ITGAL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forberg E.,
 RA Gullberg D.;
 RT "allb1 integrin is important for mesenchymal cell function:
 RT elimination of allb1 leads to dwarfism";
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johansson M., Popova S.N.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY124460; AAM62130.1; -;
 KW INTEGRIN.
 SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75B4271 CRC64;
 Query Match 10.3%; Score 122; DB 11; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 8.3e-118;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 81 GNTCKNLGRLVTVSVSEKDNRLGLSLATPKNSFLACSPKSHGCSGVVYTGMS 140
 DB 81 GNTCKNLGRLVTVSVSEKDNRLGLSLATPKNSFLACSPKSHGCSGVVYTGMS 140
 QY 141 RVNSNFRSKTVAPALQRCQTYMDIVILVLDGNSIYVVEVQHFILNLIKFFYIGPGQIQ 200
 DB 141 RVNSNFRSKTVAPALQRCQTYMDIVILVLDGNSIYVVEVQHFILNLIKFFYIGPGQIQ 200
 QY 201 VG 202

```
Db 201 VG 202

RESULT 3
Q8CE84 PRELIMINARY; PRT; 823 AA.
ID Q8CE84;
AC Q8CE84;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE MST018 homolog.
GN 4732459H24Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028821; BAC26137.1; -.
DR MGD; MGI:2442114; 4732459H24Rik.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00191; Int_alpha; 4.
SQ SEQUENCE 823 AA; 92264 MW; A330236324A0E089 CRC64;

Query Match 3.8%; Score 45; DB 11; Length 823;
Best Local Similarity 100.0%; Pred.No. 2.5e-37; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0;

QY 766 GWFTTLRVSPFVWNGCNEDEHCVDELVDARSDLPTAMEYQORVL 810
|||||
DB 401 GWFTTLRVSPFVWNGCNEDEHCVDELVDARSDLPTAMEYQORVL 445

RESULT 4
O42094 PRELIMINARY; PRT; 1171 AA.
ID O42094;
AC O42094;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ALPHAI integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archoosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RA MEDLINE=97476270; PubMed=9334246;
RX Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
RT the alphaI integrin gene.";
RL J. Biol. Chem. 272:26643-26651 (1997).
DR EMBL; AB000470; BAA23160.1; -.
DR EMBL; AB000471; BAA23161.1; -.
DR PIR; A55348; A55348.
DR HSP; P17301; 180X.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
```

```
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Query Match 1.3%; Score 15; DB 13; Length 1171;
Best Local Similarity 100.0%; Pred.No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
DB 162 DIVIVLDGNSIYPW 176

RESULT 5
Q8BM12 PRELIMINARY; PRT; 288 AA.
ID Q8BM12;
AC Q8BM12;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Integrin alpha-10 precursor homolog (Fragment).
GN ITGAL0.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK037222; BAC29761.1; -.
DR MGD; MGI:2153482; Itgal0.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF00357; Integrin_A; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
FT NON_TER
SQ SEQUENCE 288 AA; 31949 MW; 3F6200F9C9475EE5 CRC64;

Query Match 0.9%; Score 11; DB 11; Length 288;
Best Local Similarity 100.0%; Pred.No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 ICGLLLLLLLV 1150
|||||
DB 252 ICGLLLLLLLV 262

RESULT 6
Q91670 PRELIMINARY; PRT; 116 AA.
ID Q91670;
AC Q91670;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
```

```

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha-1 (Fragment).
GN GENE 4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929 (1996).
DR EMBL; U44025; AAC59878.1; -.
DR GO; GO:0003035; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR SMART; SM00191; Int_alpha; 1.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12749 MW; 2A8A5CA3E83DD07E CRC64;

Query Match 0.8%; Score 10; DB 13; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;

Qy 39 AFGGYTQQH 48
Db 39 AFGGYTQQH 48

RESULT 7
Q8XSH5 PRELIMINARY; PRT; 399 AA.
AC Q8XSH5;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 25, Last annotation update)
DE Putative transport transmembrane protein.
GN RSP0499 OR R500378.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Bottier P., Camus J.C., Catalicio L.,
RA Chandler M., Choine N., Claudel-Renard C., Cunac S., Denange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AL646079; CAD17650.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001179; FKBP_PPase.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00453; FKBP_PPase_1; 1.
DR PROSITE; PS00850; MFS; 1.
KW plasmid; Complete proteome.
SQ SEQUENCE 399 AA; 42148 MW; 096BB97E83DCA7D1 CRC64;

Query Match 0.8%; Score 10; DB 16; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1152 GLLLLALLVL 1161
Db 161 GLLLLALLVL 170

RESULT 8
Q8N112 PRELIMINARY; PRT; 164 AA.
AC Q8N112;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical gene LOC132228 (Hypothetical protein FLJ38608).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Negahari K., Masuno Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028000; AAH28000.1; -.
DR EMBL; AK095927; BAC04652.1; -.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17868 MW; 509993458714252E CRC64;

Query Match 0.8%; Score 9; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLLLALLVL 1161
Db 99 LLLLLALLVL 107

RESULT 9
Q8YCR3 PRELIMINARY; PRT; 315 AA.
AC Q8YCR3;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE PERMEASE.
GN BMEII0465.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=2002109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muejer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonowski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";

```

```

Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009683; JAL53707.1; -.
DR PIR; AH3567; AH3567.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 33486 MW; 7A34836134FB2F7D CRC64;

Query Match      0.8%; Score 9; DB 16; Length 315;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
DB 283 LLLALLVLA 291

RESULT 10
Q8FVL4 PRELIMINARY; PRT; 316 AA.
AC Q8FVL4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Membrane protein, putative.
GN BRA0823.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014576; AAN3398.1; -.
DR TIGR; BRA0823; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Complete proteome.
SQ SEQUENCE 316 AA; 33671 MW; 4DC0DC32D8AB0807 CRC64;

Query Match      0.8%; Score 9; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
DB 284 LLLALLVLA 292

RESULT 11
Q8HY16 PRELIMINARY; PRT; 400 AA.
AC Q8HY16;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Poliovirus receptor.
GN FVR.
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
OX NCBI_TaxID=9515;
RN [1]

```

```

SEQUENCE FROM N.A.
RA Ida-Hosonuma M., Sasaki Y., Toyoda H., Nomoto A., Gotoh O.,
RA Yonekawa H., Koike S.;
RT "Host range of poliovirus is restricted to simians because of a rapid
RT sequence change of the poliovirus receptor gene during evolution.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086131; BAC41711.1; -.
DR EMBL; AB086124; BAC41711.1; JOINED.
DR EMBL; AB086125; BAC41711.1; JOINED.
DR EMBL; AB086126; BAC41711.1; JOINED.
DR EMBL; AB086127; BAC41711.1; JOINED.
DR EMBL; AB086128; BAC41711.1; JOINED.
DR EMBL; AB086129; BAC41711.1; JOINED.
DR EMBL; AB086130; BAC41711.1; JOINED.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS50835; IG_LIKE; 3.
KW Receptor.
SQ SEQUENCE 400 AA; 43322 MW; C6C3A59A0A285F3A CRC64;

Query Match      0.8%; Score 9; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVLA 1161
DB 9 LLLALLVLA 17

RESULT 12
Q89LS8 PRELIMINARY; PRT; 413 AA.
AC Q89LS8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Blr4465 protein.
GN BLR4465.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484988; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005951; BAC49730.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016811; F:hydrolase activity, acting on carbon-nitrog. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR004304; FmdA_AmdA_1.
DR Pfam; PF03669; FmdA_AmdA_1.
DR PROSITE; PS00190; CYTCCHROME_C; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 45159 MW; E2F309DCC2E9D82B CRC64;

Query Match      0.8%; Score 9; DB 16; Length 413;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 148 FSKTVAPAL 156
DB 18 FSKTVAPAL 26

RESULT 13
QY 073804 PRELIMINARY; PRT; 497 AA.
AC 073804;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OC NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99177347; PubMed=10077531;
RA Gellner K., Brenner S.;
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
rubripes";
RL Genome Res. 9:251-258(1999).
DR EMBL; AF056116; AAC34385.1; -.
DR InterPro; IPR008075; Lipocalin_l_recep.
DR InterPro; IPR006876; LMBRI.
DR Pfam; PF04791; LMBRI; 1.
DR PRINTS; PR01692; LIPOCALINIMR.
KW Hypothetical protein.
SQ SEQUENCE 497 AA; 56376 MW; 802743598B83B64 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 497;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161
DB 169 LLLALLVL 177

RESULT 14
QY 039775 PRELIMINARY; PRT; 607 AA.
AC 039775;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Legumin, 11S globulin.
OS Gnetum gnemon (Bago).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales.
OC NCBI_TaxID=3382;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Seed;
RX MEDLINE=98440765; PubMed=9767693;
RA Shukov A.D., Braun H., Chesnokov Y.V., Horstmann C., Kakhovskaya I.A.,
RA Baumeister H.;
RT "Sequence peculiarity of Gnetalean legumin-like seed storage
proteins";
RL J. Mol. Evol. 47:486-492(1998).
DR EMBL; Z50779; CAA90642.1; -.
DR PIR; S60658; S60658.
DR GO; GO:004573; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
SQ SEQUENCE 607 AA; 68910 MW; 7AD6DDC604FB3A60 CRC64;

```

```

Query Match 0.8%; Score 9; DB 10; Length 607;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161
DB 10 LLLALLVL 18

RESULT 15
QY 06271 PRELIMINARY; PRT; 780 AA.
AC 06271;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha-2 subunit (fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX Meng F., Desimone D.W.;
RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE OF 95-168 FROM N.A.
RX MEDLINE=94008528; PubMed=8404528;
RA Whittaker C.A., Desimone D.W.;
RT "Integrin alpha subunit mRNAs are differentially expressed in early
Xenopus embryos";
RL Development 117:1239-1249(1993).
DR EMBL; L43058; AAA69770.1; -.
DR PIR; I51524; I51524.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SMC0191; Int_alpha; 4.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR NON TER 1
FT NON TER 1
SQ SEQUENCE 780 AA; 87017 MW; 9518B18C2B6BF637 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 780;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGP 503
DB 108 TDVLLVGP 116

RESULT 16
QY 09TS65 PRELIMINARY; PRT; 54 AA.
AC 09TS65;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE LEUKOINTEGRIN alpha D beta 2 (Fragments).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95325609; PubMed=7541420;
RX Danilenko D.M., Rossitto P.V., Van der Vieren M., Le Trong H.,

```

```

RA McDonough S.P., Affolter V.K., Moore P.F.;
RT "A novel canine leukointegrin, alpha d beta 2, is expressed by
RT specific macrophage subpopulations in tissue and a minor CD8+
RT lymphocyte subpopulation in peripheral blood.";
PL J. Immunol. 155:35-44(1995).
FT NON_TER 1
FT NON_CONS 23
FT NON_TER 54
SQ SEQUENCE 54 AA; 5624 MW; C2FC92C2FDEC3E6C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 54;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 24 LVVGAPLE 31

RESULT 17
Q88NW3 ID Q88NW3 PRELIMINARY; PRT; 58 AA.
AC Q88NW3;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN Pseudomonas putida (strain KT2440).
OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2243060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzaple E., Scanlan D., Tran K., Moazzez A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eissen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016778; AAN66716.1; -.
DR TIGR; PP1091; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6734 MW; CD9350CEF793D269 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 EERLQKQ 891
Db 43 EERLQKQ 50

RESULT 18
Q8CEU3 ID Q8CEU3 PRELIMINARY; PRT; 76 AA.
AC Q8CEU3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE CD24a antigen.
GN CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK014279; BAC25429.1; -.
DR PIR; PT0721; PT0721.
DR MGD; MGI:88323; Cd24a.
SQ SEQUENCE 76 AA; 7787 MW; 6853F12240F91AEB CRC64;

Query Match 0.7%; Score 8; DB 11; Length 76;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 19
Q9N241 ID Q9N241 PRELIMINARY; PRT; 78 AA.
AC Q9N241;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Saimiri boliviensis (Bolivian squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=27679;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seanan M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125670; AAF64385.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 78
SQ SEQUENCE 78 AA; 7607 MW; D4E684AD6F6375B CRC64;

Query Match 0.7%; Score 8; DB 6; Length 78;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 64 LLLALLVL 71

RESULT 20
Q9N242 ID Q9N242 PRELIMINARY; PRT; 87 AA.
AC Q9N242;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

```

```

DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Ateles sp. (Spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=9311;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20215006; PubMed=10750051;
RX Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125669; AAF64384.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 87 AA; 8388 MW; 64574359DAC555D4 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 73 LLLALLVL 80
|||||

RESULT 21
Q99587 PRELIMINARY; PRT; 88 AA.
AC Q99587;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE Dopamine D4 receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96362071; PubMed=8746407;
RX Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,
RX Lichtermann D., Minges J., Albus J., Bormann M., Franzek E., et al;
RT "Identification of two novel polymorphisms and a rare deletion variant
RT in the human dopamine D4 receptor gene.";
RL Psychiatr. Genet. 5:97-103 (1995).
DR EMBL; S82918; AAB46803.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 88 AA; 8482 MW; 314093AB4678CEDF CRC64;

Query Match 0.7%; Score 8; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161

```

```

DB 74 LLLALLVL 81
|||||

RESULT 22
Q9N244 PRELIMINARY; PRT; 91 AA.
AC Q9N244;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Nasalis larvatus (Proboscis monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Nasalis.
OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20215006; PubMed=10750051;
RX Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125667; AAF64382.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 91 AA; 8827 MW; 67B53D0764BD439C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84
|||||

RESULT 23
Q9N245 PRELIMINARY; PRT; 91 AA.
AC Q9N245;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20215006; PubMed=10750051;
RX Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125666; AAF64381.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.

```



```
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8728 MW; 7CFF8CAC7E0CF39C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 77 LLLALLVL 84

RESULT 24
ID O97517 PRELIMINARY; PRT; 91 AA.
AC O97517;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010297; AAC67225.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004872; F:receptor activity; IEA.
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8851 MW; 9DF2F56F1D7CF828 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 77 LLLALLVL 84

RESULT 25
O97515 PRELIMINARY; PRT; 91 AA.
AC O97515;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9597;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010295; AAC67223.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004872; F:receptor activity; IEA.
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8763 MW; 76BDBF6BE38E299 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 77 LLLALLVL 84

RESULT 26
O97521 PRELIMINARY; PRT; 91 AA.
AC O97521;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010301; AAC67229.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004872; F:receptor activity; IEA.
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8934 MW; 93CD0F1CFC8E82C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 77 LLLALLVL 84

RESULT 27
O9N243 PRELIMINARY; PRT; 91 AA.
ID O9N243
```

```

AC Q9N243;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Presbytis cristata (Silvered langur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae;
OC Presbytis.
OX NCBI_TaxID=36232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125668; AA664383.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 91
FT SEQUENCE 91 AA; 8827 MW; 67E53D0764BD439C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84
|||||

RESULT 28
Q9N246 PRELIMINARY; PRT; 91 AA.
AC Q9N246;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125665; AA64380.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 91
FT SEQUENCE 91 AA; 8770 MW; 608F90C00E10839C CRC64;

```

```

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84
|||||

RESULT 29
Q9N247 PRELIMINARY; PRT; 93 AA.
AC Q9N247;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9590;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125664; AA64379.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 93
FT SEQUENCE 93 AA; 8904 MW; 5D07B793F2B624E4 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 93;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 79 LLLALLVL 86
|||||

RESULT 30
Q99586 PRELIMINARY; PRT; 95 AA.
AC Q99586;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96382071; PubMed=8746407;
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,
RA Lichtermann D., Minges J., Albus J., Bormann M., Franzek E., et al;
RT "Identification of two novel polymorphisms and a rare deletion variant
RT in the human dopamine D4 receptor gene.";
RL Psychiatr. Genet. 5:97-103(1995).
DR EMBL: S82917; AAB46802.1; -.

```

DR GO; GO:0004952; F:dopamine receptor activity; NAS.
DR GO; GO:0007212; P:dopamine receptor signaling pathway; NAS.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9148 MW; 7163AF13B6156686 CRC64;

Query Match 0.7%; Score 8; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 31
ID O97514 PRELIMINARY; PRT; 95 AA.
AC O97514;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010294; AAC67222.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 32
ID O9N249 PRELIMINARY; PRT; 95 AA.
AC O9N249;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates klossii (Kloss's gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 33
ID O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34
ID O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34

OX NCBI_TaxID=9597;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125662; AAF64377.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 33
ID O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34
ID O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34

```

O97516
ID O97516 PRELIMINARY; PRT; 95 AA.
AC O97516;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010296; AAC67224.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PSS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 95 AA; 9121 MW; ECF6C50B3A809D33 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 35
O97520
ID O97520 PRELIMINARY; PRT; 95 AA.
AC O97520;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010296; AAC67224.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PSS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 95 AA; 9221 MW; 96FB63A12B8A3647 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 35
O97520
ID O97520 PRELIMINARY; PRT; 95 AA.
AC O97520;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010296; AAC67224.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PSS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 95 AA; 9221 MW; 96FB63A12B8A3647 CRC64;

```

```

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 36
O97519
ID O97519 PRELIMINARY; PRT; 95 AA.
AC O97519;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010299; AAC67227.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PSS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 95 AA; 9221 MW; 96FB63A12B8A3647 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 37
O97518
ID O97518 PRELIMINARY; PRT; 99 AA.
AC O97518;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010298; AAC67226.1; -.

```

```
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:004872; F: receptor activity; IEA.
DR GO: GO:001584; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin...; IEA.
DR InterPro: IPR000216; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
FT NON_TER 99 99
SQ SEQUENCE 99 AA; 9335 MW; D4CED9750184BC47 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 85 LLLALLVL 92

RESULT 38
Q8E9P1 PRELIMINARY; PRT; 104 AA.
AC Q8E9P1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cell division protein FtsL.
GN FtsL OR S04226.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=MR-1;
MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Taapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015855; AAN57198.1; -.
DR TIGR; SO4226; -.
DR InterPro; IPR007082; FtsL.
DR Pfam; PF04999; FtsL; 1.
DR Complete proteome.
SQ SEQUENCE 104 AA; 12187 MW; 408B210954032920 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 23 LLLALLVL 30

RESULT 39
Q9SDS4 PRELIMINARY; PRT; 106 AA.
AC Q9SDS4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Non-specific lipid transfer protein.

GN LPTI.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=cv._Hanbyul; TISSUE=Leaf;
RA Hwang B.K., Jung H.W.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208832; AAF23458.1; -.
DR HSSP; P23096; 1RZL.
DR GO: GO:0008289; F: lipid binding; IEA.
DR GO: GO:0006889; P: lipid transport; IEA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryptophan amyl; 1.
DR PRINTS; PRC0382; LIPIDTRANSFER.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 106 AA; 11291 MW; 5E9C43238B9381C9 CRC64;

Query Match 0.7%; Score 8; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GOIQGVV 204
DB 20 GOIQGVV 27

RESULT 40
Q8R1E9 PRELIMINARY; PRT; 112 AA.
AC Q8R1E9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE FROM N.A.
TISSUE=Salivary gland;
RA Strausberg R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024677; AAH24677.1; -.
DR MGD; MGI:2655426; C2a.
DR InterPro; IPR006038; Uteroglobin_supf.
DR Hypothetical protein.
SQ SEQUENCE 112 AA; 12371 MW; 69CE957825183F72 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1160
DB 5 LLLALLVL 12

RESULT 41
Q8JZX1 PRELIMINARY; PRT; 112 AA.
AC Q8JZX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Salivary androgen-binding protein gamma
DE subunit).
```

C2D OR ABPG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1) LLLLLL
 RN (2) LLLLLL
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN (1) LLLLLL
 RN (2) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 RN (3) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on
 Chromosome 7: Characterization and evolutionary relationships.";
 RL Mamm. Genome 0:0-0(2003).
 DR EMBL; BC035222; AAH35222.1; -.
 DR EMBL; AY325898; AAP94990.1; -.
 DR MGD; MGI:2655745; C2d.
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 12794 MW; F83285868F4BD53D CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLL 1160
 DB 5 LLLLLL 12

RESULT 42

Q7TNN5 PRELIMINARY; PRT; 112 AA.
 AC Q7TNN5;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein beta subunit.
 GN ABPG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1) LLLLLL
 RN (2) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 RN (3) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on
 Chromosome 7: Characterization and evolutionary relationships.";
 RL Mamm. Genome 0:0-0(2003).
 DR EMBL; AY325897; AAP94989.1; -.
 SQ SEQUENCE 112 AA; 12383 MW; EB5F2C68DBF8D17C CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLL 1160
 DB 5 LLLLLL 12

RESULT 43

Q7TNN0 PRELIMINARY; PRT; 112 AA.
 AC Q7TNN0;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein beta subunit.
 GN ABPG.
 OS Mus musculus musculus (eastern European house mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=39442;
 RN (1) LLLLLL
 RN (2) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 DR EMBL; AY293278; AAP44465.1; -.
 SQ SEQUENCE 112 AA; 12411 MW; 818431A8CEFF8C029 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLL 1160
 DB 5 LLLLLL 12

RESULT 44

Q7TNN7 PRELIMINARY; PRT; 112 AA.
 AC Q7TNN7;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein gamma subunit.
 GN ABPG.
 OS Mus musculus musculus (eastern European house mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=39442;
 RN (1) LLLLLL
 RN (2) LLLLLL
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 DR EMBL; AY293281; AAP44468.1; -.
 SQ SEQUENCE 112 AA; 12546 MW; 9F955317D9116883 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLL 1160

```

Db      TIGR; DR2210; -
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 127 AA; 14314 MW; 2B3679D3522A0B1 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred. No. 27;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1152 GLLLLALL 1159
DB      7 GLLLLALL 14

RESULT 47
Q82U09      PRELIMINARY;      PRT;      127 AA.
AC      Q82U09
DT      01-JUN-2003 (TREMBlrel. 24, Created)
DT      01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      CRCB protein.
GN      NEI1704.
OS      Nitrosomonas europaea.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC      Nitrosomonadaceae; Nitrosomonas.
OX      NCBI_TaxID=915;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 19718 / IF0 14298;
RX      MEDLINE=22586410; PubMed=12700255;
RA      Chain P., Laverdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA      Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA      Arciero D.M., Hommes N.G., Whittaker M.M., Ayr D.J.,
RT      "Complete genome sequence of the ammonia-oxidizing bacterium and
RT      obligate chemolithoautotroph Nitrosomonas europaea.";
RL      J. Bacteriol. 185:2759-2773 (2003).
DR      EMBL; BX321862; CAD85615.1; -
DR      GO; GO:0016020; C:membrane; IEA.
DR      InterPro; IPR003691; Camphor_Crb.
DR      Pfam; PF02537; CRCB; 1.
DR      TIGRFAMs; TIGR00494; crCB; 1.
KW      Complete proteome.
SQ      SEQUENCE 127 AA; 13419 MW; 55F30C8DD8AD7952 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred. No. 27;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1147 GSTLGLL 1154
DB      11 GSTLGLL 18

RESULT 48
Q8C442      PRELIMINARY;      PRT;      137 AA.
AC      Q8C442
DT      01-MAR-2003 (TREMBlrel. 23, Created)
DT      01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT      01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE      Stimulated by retinoic acid gene 6.
GN      STRA6.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Hippocampus;
RC      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium.
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of

```

```

RT 60,770 full-length cDNAs.;
RL Nature 420:563-573(2002).
DR EMBL; AK083120; BAC38769.1; -.
DR PIR; PT0714; PT0714.
DR MGI; MGI:107742; Stra6.
SQ SEQUENCE 137 AA; 14573 MW; 27A54C974EF9670C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 11; Length 137;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 LLLALLLV 1160
DB 64 LLLALLLV 71

RESULT 49
Q61692 PRELIMINARY; PRT; 141 AA.
ID Q61692
AC Q61692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative heat stable antigen.
GN HSA-C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA X C57BL/6; TISSUE=Spleen;
RX MEDLINE=91209380; PubMed=2019286;
RA Wenger R.H., Ayane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called the
RT heat-stable antigen.";
RL Eur. J. Immunol. 21:1039-1046(1991).
DR EMBL; X56486; CAA3843.1; -.
DR PIR; S15785; S15785.
SQ SEQUENCE 141 AA; 15515 MW; E4BF6428ADC03C69 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 11; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 12 GLLLLALL 19

RESULT 50
Q71LM3 PRELIMINARY; PRT; 144 AA.
ID Q71LM3
AC Q71LM3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bradykinin-potentiating/c-type natriuretic peptide isoform 2
DE (Fragment)
OS Bothrops jararacussu (Jararacussu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8726;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Lameu C., Radis-Baptista G., Yamane T.,
RA Camargo A.C.M.;
RT "Cloning and sequence analysis of a Bothrops jararacussu Bpps
RT precursor.";
RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY310915; AAP83422.1; -.

FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15982 MW; B23FED5230A84910 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 13; Length 144;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 51
Q8QG90 PRELIMINARY; PRT; 146 AA.
ID Q8QG90
AC Q8QG90;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bradykinin-potentiating protein.
OS Bothrops insularis (Island jararaca) (Queimada jararaca)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490532; AAM09691.1; -.
SQ SEQUENCE 146 AA; 16088 MW; F58244C4F24B1799 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 13; Length 146;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 52
Q7XK40 PRELIMINARY; PRT; 161 AA.
ID Q7XK40
AC Q7XK40;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNBA004K18.5 protein.
GN OSJNBA004K18.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606595; CAE05863.1; -.
SQ SEQUENCE 161 AA; 16300 MW; 0E8B2DF9C58ADC61 CRC64;

```



```

Query Match          0.7%; Score 8; DB 10; Length 161;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 LALLVLAL 1163
DB 136 LALLVLAL 143

RESULT 53
Q98MM8 PRELIMINARY; PRT; 174 AA.
AC Q98MM8;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-NAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein mlr0512.
GN MLR0512.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP002995; BAB48085.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 174 AA; 20270 MW; E93E12118C11022F CRC64;

Query Match          0.7%; Score 8; DB 16; Length 174;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 LRSLLKALK 1097
DB 38 LRSLLKALK 45

RESULT 54
P79786 PRELIMINARY; PRT; 181 AA.
AC P79786;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Thrombospondin-1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97209337; PubMed=9056637;
RA Tucker R.P., Haglos C., Chiquet-Ehrismann R., Lawler J.;
RT "In situ localization of thrombospondin-1 and thrombospondin-3
RT transcripts in the avian embryo.";
RL Dev. Dyn. 208:326-337 (1997).
DR EMBL; U76984; AAB19208.1; -;
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003367; tsp_3.
DR Pfam; PF02412; tsp_3; 8.

Query Match          0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 VDIDGDGV 494
DB 116 VDIDGDGV 123

RESULT 55
Q90Y12 PRELIMINARY; PRT; 181 AA.
AC Q90Y12;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 1.
OS Crotales durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;
RT "Crotales durissus terrificus bradykinin potentiating peptide
RT precursor";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308593; AAL09426.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP.1.
DR PROSITE; PS00263; NATRIURITIC_PEPTIDE; 1.
SQ SEQUENCE 181 AA; 18560 MW; 7B5ADC5B9372D07F CRC64;

Query Match          0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
AC Q90Y11;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotales durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;

```

RT "Crotalus durissus terrificus bradykinin-potentiating peptide and C-
 RL type natriuretic peptide precursor isoform2.";
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF308594; AAL09427.1; -
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005179; F:hormone activity; IEA.
 DR InterPro; IPR000663; Natri_peptide.

DR Pfam; PF00212; ANP; 1.
 DR PRINTS; PRO0710; NATPEPTIDES.
 DR SMART; SM00183; NAT_PEP; 1.
 DR PROSITE; PS00263; Natriuretic_Peptide; 1.

SQ SEQUENCE 181 AA; 18507 MW; 982B95D38A5FF27 CRC64;
 Query Match 0.7%; Score 8; DB 13; Length 181;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
 DB 10 GLLLLALL 17

RESULT 57

Q8U4N4 PRELIMINARY; PRT; 222 AA.
 ID Q8U4N4
 AC Q8U4N4

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF0047.

GN PF0047.
 OS Pyrococcus furiosus
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]

RP STRAIN=FROM N.A.
 RC STRAIN=V61 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;

RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF010131; AAL080171.1; -
 DR Hypothetical protein; Complete proteome;

KW Pyrococcus furiosus; Complete proteome;
 SQ SEQUENCE 222 AA; 24556 MW; 721D79ED547A0F13 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 222;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGKVVYVE 519
 DB 106 RGKVVYVE 113

RESULT 58

Q7V5L4 PRELIMINARY; PRT; 224 AA.
 ID Q7V5L4
 AC Q7V5L4

DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Conserved hypothetical membrane protein.

GN F3H9.20 OR PWT1539.
 OS Prochlorococcus marinus (strain MIT 9313).

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.

OX NCBI_TaxID=74547;
 RN [1]

RP SEQUENCE FROM N.A.
 RC MEDLINE=22825698; PubMed=12917642;
 RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,

RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,

RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "Genome divergence in two prochlorococcus ecotypes reflects oceanic
 RT niche differentiation.";
 RL Nature 424:1042-1047(2003).
 DR EMBL; BX572099; CAE21714.1; -
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 224 AA; 24172 MW; E1CD9F149C279A29 CRC64;
 Query Match 0.7%; Score 8; DB 16; Length 224;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 LLVLA LRK 1165
 DB 190 LLVLA LRK 197

RESULT 59

O57835 PRELIMINARY; PRT; 224 AA.
 ID O57835
 AC O57835

DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PH0087.

GN PH0087.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]

RP STRAIN=FROM N.A.
 RC STRAIN=OT3;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Koeugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.
 RT "Complete sequence and gene organization of the genome of a hyper-

RT thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).

DR EMBL; AP000001; BAA29156.1; -
 DR PIR; E71228; E71228.

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 224 AA; 25134 MW; 9CC56212064558E0 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 224;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGKVVYVE 519
 DB 105 RGKVVYVE 112

RESULT 60

P74604 PRELIMINARY; PRT; 226 AA.
 ID P74604
 AC P74604

DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein slr1571.

GN SLR1571.
 OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OC NCBI_TaxID=1148;
 RN [1]

RP SEQUENCE FROM N.A.
 RC MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirasawa M., Sugura M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shingo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90916; BAA18712.1; -;
 DR FIR; S76800; S76800.
 DR InterPro; IPR002838; DUF124.
 DR Pfam; PF01987; DUF124; 1.
 DR ProDom; PD013634; DUF124; 1.
 DR TIGRPFMs; TIGR00266; TIGR00266; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 226 AA; 25003 MW; 29C5C137C128DAC0 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 226;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1073 RLVPNQE1 1080
 DB 183 RLVPNQE1 190

RESULT 61
 ID Q882P2 PRELIMINARY; PRT; 227 AA.
 AC Q882P2;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN PSPTQ2584.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Coardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brikkac L., Beaman M., Haft D., Selengut J., Nelson W., Davidson T.,
 RA White C., Fraser C., Collier A.;
 RT "Complete sequence of *Pseudomonas syringae*.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016865; AAC56088.1; -;
 DR TIGR; PSPTQ2584; -;
 DR InterPro; IPR002489; DUF14.
 DR Pfam; PF01493; GXGXG; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 227 AA; 24036 MW; AB3459F347BBF145 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 227;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 547 ASVRDLNQ 554
 DB 9 ASVRDLNQ 16

RESULT 62
 ID Q9AJX4 PRELIMINARY; PRT; 231 AA.
 AC Q9AJX4;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE Putative oxidoreductase.
 GN SC01803 OR SCI33.02.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Harris D.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC EMBL; AL939110; CAC28543.1; -;
 DR HSP; 064105; 10AA.
 DR GO; GO:0016481; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 231 AA; 23789 MW; 6AEDB40306B94086 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 231;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 790 DLVLDRS 797
 DB 27 DLVLDRS 34

RESULT 63
 ID Q98GF8 PRELIMINARY; PRT; 235 AA.
 AC Q98GF8;
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE ATP-binding component of phosphonate transport, HisP-like
 DE nucleotide binding protein, PhnL.
 GN MLR3349.

OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 ON NCBI_TaxID=381;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RA MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003001; BAB50258.1; --
 DR GO; GO:0015020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00005; ABC trans.; 1.
 DR ProDom; PD000006; ABC transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 235 AA; 25098 MW; 6FA8F03F827361 CRC64;
 Query Match 0.7%; Score 8; DB 16; Length 235;
 Best Local Similarity 100.0%; Pred.No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 EDGLIDLA 623
 DB 73 EDGLIDLA 80
 RESULT 64
 ID Q20220 PRELIMINARY; PRT; 245 AA.
 AC Q20220;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein F40F4.7.
 GN F40F4.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Felodierinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RT "The sequence of C. elegans cosmid F40F4.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; U40420; AAK84526.3; --
 DR WormPep; F40F4.7; CE28316.
 DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
 DR InterPro; IPR000182; GCSacetyl_trans.
 DR Pfam; PF00583; Acetyltransf; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 245 AA; 27893 MW; A53289FBB2F91E75 CRC64;
 Query Match 0.7%; Score 8; DB 5; Length 245;
 Best Local Similarity 100.0%; Pred.No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 YNDVVVGA 564
 DB 141 YNDVVVGA 148
 RESULT 65
 ID Q8Y1R9 PRELIMINARY; PRT; 257 AA.
 AC Q8Y1R9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein RSC0620.
 GN RSC0620 OR RS01521.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 ON NCBI_TaxID=305;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMT1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chardier M., Choisme N., Claudel-Renard C., Cunha S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:457-502(2002).
 DR EMBL; AL646060; CAD1450.1; --
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000572; Oxidored_molyb.
 DR Pfam; PF00174; oxidored_molyb; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 257 AA; 29260 MW; DAE569EB1572113D CRC64;
 Query Match 0.7%; Score 8; DB 16; Length 257;
 Best Local Similarity 100.0%; Pred.No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1149 TLGGLLLL 1156
 DB 37 TLGGLLLL 44
 RESULT 66
 ID Q8QG91 PRELIMINARY; PRT; 265 AA.
 AC Q8QG91;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Bradykinin-potentiating/c-type natriuretic protein.
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 ON NCBI_TaxID=8723;
 RX [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs)."; the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490531; AAM09690.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 0EAE1408B42358BE CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred.No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 67
Q9PW56 PRELIMINARY; PRT; 265 AA.
AC Q9PW56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide.
OS Bothrops jararaca (Jarakaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Hayashi M.A.F., Murbach A.F., Camargo A.C.M.;
RT "The precursor of C-type natriuretic peptide of snake brain contains
RT angiotensin converting enzyme inhibitors, specific for the C-catalytic
RT site".
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF171670; AAD51326.2; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 8E99AEC976CCD439 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred.No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 68
Q861M5 PRELIMINARY; PRT; 272 AA.
AC Q861M5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).

OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=107;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176095; AAO18112.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30576 MW; 8EC9E0C183D319FD CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;
Best Local Similarity 100.0%; Pred.No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
Db 232 VLLGAVGA 239

RESULT 69
Q861K3 PRELIMINARY; PRT; 272 AA.
AC Q861K3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176118; AAO18134.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30220 MW; F7FC66FFB06558CF CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;

```

Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
DB 232 VLLGAVGA 239
|||||

RESULT 70
Q8P8NG PRELIMINARY; PRT; 278 AA.
AC Q8P8NG;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE ABC transporter sugar permease.
GN LAGC OR XCC2204.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Caracotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fortighieri B.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.P.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RA host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012328; AAM41484.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; P:transp.; IEA.
DR Pfam; PF00528; BPD_transp; 1.
KW Complete proteome.
SQ SEQUENCE 278 AA; 30673 MW; FB45244D924B7361 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1159
DB 17 GGLLLAL 24
|||||

RESULT 71
Q8TPV4 PRELIMINARY; PRT; 290 AA.
AC Q8TPV4;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Uncharacterized domain specific for M.kandleri, MK-34 family, a.
GN MK1344.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.

```

```

OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
DR EMBL; AE010427; AAM02557.1; --
KW Complete proteome.
SQ SEQUENCE 290 AA; 32754 MW; 706E6883976ADF91 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 290;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VITDGESH 275
DB 161 VITDGESH 168
|||||

RESULT 72
Q94K03 PRELIMINARY; PRT; 299 AA.
AC Q94K03;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Unknown protein.
GN A:IG002N01.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370497; AAK43874.1; --
KW Complete proteome.
SQ SEQUENCE 299 AA; 34028 MW; BB2AB03C4A184D3E CRC64;

Query Match 0.7%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 44 LLLALLLV 51
|||||

RESULT 73
Q98BD8 PRELIMINARY; PRT; 307 AA.
AC Q98BD8;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein mll5620.
GN MLL5620.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobiaceae.
OX NCBI_TaxID=381;

```

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=NAFF303059; PubMed=11214968;
RA  Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA  Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Klmura T.,
RA  Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA  Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA  Takeuchi C., Yamada M., Tabata S.;
RT  "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT  Mesorhizobium loti.";
RL  DNA Res. 7:331-338(2000).
DR  ENBL; AP003007; BAB52034.1; -.
DR  InterPro; IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 307 AA; 31951 MW; 938PFCID3BC4AC16 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 307;
Best Local Similarity 100.0%; Pred.No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 74
Q92S85 PRELIMINARY; PRT; 308 AA.
AC  Q92S85;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical transmembrane protein SMC02232.
GN  R00537 OR SMC02232.
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX  NCBI_TaxID=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RA  Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA  Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA  Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA  Pohl T., Portetalle D., Puhler A., Purnelle B., Ranepberger U.,
RA  Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT  "Analysis of the chromosome sequence of the legume symbiont
RT  Sinorhizobium meliloti strain 1021.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR  EMBL; AL591784; CAC45109.1; -.
DR  InterPro; IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 308 AA; 32465 MW; CB6497E6E16BC0D CRC64;

Query Match      0.7%; Score 8; DB 16; Length 308;
Best Local Similarity 100.0%; Pred.No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 75
Q8CB84 PRELIMINARY; PRT; 323 AA.
ID  Q8CB84;
AC  Q8CB84;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```

DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Integrin alpha 2 (Fragment).
GN  ITGA2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Bone;
RC  MEDLINE=22354683; PubMed=12466851;
RA  The FANTOM Consortium,
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RT  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
DR  EMBL; AK036572; BAC29485.1; -.
DR  MGD; MGI:96600; Itga2.
DR  GO; GO:0008305; C:integrin complex; IEA.
DR  GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR  GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR  InterPro; IPR000413; Integrin_alpha.
DR  InterPro; IPR002035; VWF_A.
DR  Pfam; PF00092; vwa; 1.
DR  PRINTS; PR00453; VWFADOMAIN.
DR  SMART; SM00191; Int_alpha; 1.
DR  SMART; SM00327; VWA; 1.
DR  PROSITE; PS50234; VWFA; 1.
FT  NON_TER 323
FT  SEQUENCE 323 AA; 35025 MW; 05A4E6A44283724A CRC64;

Query Match      0.7%; Score 8; DB 11; Length 323;
Best Local Similarity 100.0%; Pred.No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  71 GDVYKCPV 78
DB  75 GDVYKCPV 82
|||||

Search completed: September 21, 2004, 13:41:59
Job time : 78 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:05:19 ; Search time 148 Seconds
(without alignments)
2268.016 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPGLVVAWALSMPGFT.....FRSARRREPGLDTPPKVLE 1188

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 262547505 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	4	AAB30929 Amino aci
2	928	78.1	1189	4	ABG12949 Novel hum
3	926	77.9	1188	4	AAU14231 Human nov
4	926	77.9	1188	4	AAB50085 Human A25
5	926	77.9	1188	5	AAU10551 Human A25
6	911	76.7	1188	4	AAU14467 Human nov
7	911	76.7	1188	7	ADE09956 Novel pro
8	911	76.7	1189	3	AAB25582 IRGALL pr
9	911	76.7	1189	6	ABR58364 Human nov
10	911	76.7	1189	6	ADA27054 Human nov
11	911	76.7	1189	7	ADE63570 Human Pro
12	911	76.7	1189	8	ADE6584 Novel hum
13	810	68.2	1034	3	AAB25590 Protein e
14	810	68.2	1034	6	ADA27062 Human nov
15	810	68.2	1034	8	ADE6592 Novel hum
16	753	63.4	1120	6	ABR58365 Human nov
17	489	41.2	707	4	AAU19663 Human nov
18	489	41.2	707	5	ABP47883 Human pol
19	489	41.2	707	7	ADC10845 Human ext
20	193	16.2	193	5	AAU76854 Human int
21	190	16.0	193	5	AAU76863 Human int
22	122	10.3	1188	4	AAB50087 Murine A2
23	122	10.3	1188	5	AAU10552 Murine A2
24	104	8.8	360	7	ADE08585 Novel pro
25	89	7.5	109	5	ABG66673 Human nov

26	88	7.4	103	4	AAU19822 Human nov
27	88	7.4	103	4	AAU87675 Novel cen
28	88	7.4	103	5	ABP48042 Human pol
29	88	7.4	103	7	ADC11004 Human pro
30	42	3.5	42	7	ABD32059 alphan-in
31	42	3.5	545	5	ABP72288 Murine pr
32	42	3.5	688	5	ABB72300 Rat prote
33	42	3.5	696	5	ABB72289 Rat prote
34	38	3.2	757	4	AAU39928 Human pol
35	34	2.9	117	4	ABG12950 Novel hum
36	33	2.8	33	3	AAB25614 Human sec
37	33	2.8	33	3	ADA27114 Human nov
38	33	2.8	33	8	ADE86651 Transmemb
39	24	2.0	24	4	AAU30927 Peptide d
40	21	1.8	158	3	AAU76112 Rat integ
41	21	1.8	158	4	AAB56051 Skin cell
42	21	1.8	158	5	ABB72251 Rat prote
43	17	1.4	85	3	AAU76111 Rat integ
44	17	1.4	85	4	AAB56050 Skin cell
45	17	1.4	85	5	ABB72250 Rat prote
46	15	1.3	148	4	AAU19634 Human nov
47	15	1.3	148	4	AAU19794 Human nov
48	15	1.3	148	5	ABP47854 Human pol
49	15	1.3	148	5	ABP48014 Human pol
50	15	1.3	148	7	ADC10816 Human ext
51	15	1.3	148	7	ADC10976 Human pro
52	15	1.3	192	6	AAE33535 Rat-human
53	15	1.3	192	6	AAE33536 Rat alpha
54	15	1.3	192	6	AAE33537 Human alp
55	15	1.3	195	5	AAU76851 Human int
56	15	1.3	195	5	AAU76860 Human int
57	15	1.3	214	4	AAB50041 Rat alpha
58	15	1.3	214	4	AAB50042 Human alp
59	15	1.3	214	6	AAE33539 Rat alpha
60	15	1.3	214	6	AAE33540 Human alp
61	15	1.3	1151	8	ADE86652 Human int
62	15	1.3	1179	5	ABB90759 Human pro
63	15	1.3	1179	5	ABP64915 Human tum
64	15	1.3	1179	6	ABU54466 Human tum
65	15	1.3	1180	5	ABB90788 Rat Tumou
66	15	1.3	1180	6	ABU54495 Mouse tum
67	15	1.3	1180	7	ADE63568 Rat Prote
68	14	1.2	15	4	AAB30928 Antigenic
69	14	1.2	1183	2	AAU07728 Armenian
70	13	1.1	195	5	AAU76862 Human int
71	13	1.1	195	5	AAU76853 Human int
72	13	1.1	1132	3	AAU32243 Human int
73	13	1.1	1152	4	AAB64657 Human sec
74	13	1.1	1152	4	AAB64658 Human sec
75	13	1.1	1167	3	AAU32242 Human int
76	13	1.1	1167	4	AAB64584 Human sec
77	13	1.1	1167	6	ABP99490 Human sec
78	13	1.1	1167	6	ABR00964 Human gen
79	13	1.1	1167	6	ADA44026 Human sec
80	11	0.9	22	2	AAU12588 Platelet
81	11	0.9	43	7	ADB32060 alphaA-in
82	11	0.9	303	3	AAU32282 Mouse int
83	11	0.9	979	4	ABG29239 Novel hum
84	11	0.9	1147	7	ADD49069 Human NOV
85	11	0.9	1181	6	ABU03548 Angiogene
86	11	0.9	1181	6	ABR59703 Human ViA
87	11	0.9	1181	6	ABU03616 Human exp
88	11	0.9	1181	6	ABU03614 Human exp
89	11	0.9	1181	6	ABU03617 Armenian
90	11	0.9	1181	6	AAU07729 Integrin
91	11	0.9	1367	2	AAW70542 Human exp
92	11	0.9	1367	2	AAW70542 Human exp
93	11	0.9	1367	2	AAW70542 Human exp
94	9	0.8	9	2	AAU03615 CBS1 moti
95	8	0.7	16	3	AAU03615 CBS1 moti
96	8	0.7	17	3	AAU03615 CBS1 moti
97	8	0.7	17	3	AAU03615 CBS1 moti
98	8	0.7	17	3	AAU03615 CBS1 moti

99 8 0.7 17 2 AAW65098
100 8 0.7 17 2 AAW72832

Aaw65098 Canine be
Aaw72832 Canine al

ALIGNMENTS

RESULT 1

AAB30929
ID AAB30929 standard; protein; 1188 AA.

XX AAB30929;

XX DT 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha1 integrin chain.

DE Human; integrin; alpha1 subunit; fibroblast; muscle cell; chondrocyte;
XX osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
KW osteoporosis; cartilage damage; bone damage; cartilage.

XX Homo sapiens.

XX Key Location/Qualifiers
FT Peptide 1..22
FT /note= "signal peptide"
FT Region 951..972
FT /note= "leucine zipper"
FT Domain 1142..1164
FT /note= "transmembrane domain"

XX W0200075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SF001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX N-PSDB; AAC86871.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Disclosure; Fig 2a-c; 79pp; English.

XX The present sequence represents a human integrin subunit, designated
CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers
CC of cell target molecules, such as fibroblasts, muscle cells,
CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.
CC They are also used for determining the differential-stage of cells during
CC differentiation, development in pathological conditions, in tissue
CC regeneration, in transplantation or in therapeutic and physiological
CC repair of tissues. The pathological conditions involving subunit alpha11
CC are selected from damage of cells, muscle dystrophy, fibrosis, wound
CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,
CC damage of cartilage and bone, and cartilage and bone diseases. The
CC polypeptide is useful for detecting the formation of cartilage during
CC embryonic development, for detecting physiological therapeutic repair of
CC cartilage and muscle, for selection and analysis, or for sorting,
CC isolating or purification of chondrocytes and muscle cells, for detecting
CC regeneration of cartilage or chondrocytes during transplantation of
CC cartilage or chondrocytes during transplantation of cartilage or
CC chondrocytes, respectively, or of muscle or muscle cells during
CC transplantation of muscle or muscle cells, respectively, and for studies
CC of differentiation of chondrocytes or muscle cells

XX Sequence 1188 AA;

Query Match 100.0%; Score 1188; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDLPRGLVAVWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFGYTVQCHDISGNKWLVGGA	60
DB	1	MDLPRGLVAVWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFGYTVQCHDISGNKWLVGGA	60
QY	61	PLETNGYQKTGDVYKCPVIHGNCCTKLNLRVTLSNVSEKDNMRGLGLSLATNPKDNSFLA	120
DB	61	PLETNGYQKTGDVYKCPVIHGNCCTKLNLRVTLSNVSEKDNMRGLGLSLATNPKDNSFLA	120
QY	121	CSPLWSHECGSSYYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE	180
DB	121	CSPLWSHECGSSYYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE	180
QY	181	VQHELINILKKFYIGPQIQGVVQVQGEDVVEHFLNDYRSVDVVEAAASHIEQRCGTET	240
DB	181	VQHELINILKKFYIGPQIQGVVQVQGEDVVEHFLNDYRSVDVVEAAASHIEQRCGTET	240
QY	241	RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQSERDNVTRYAVAVL	300
DB	241	RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQSERDNVTRYAVAVL	300
QY	301	GYNRRGINPETFLNEIKYIASPDQKHFNVTDEALKDIDVALGDRIFSLEGTNNKET	360
DB	301	GYNRRGINPETFLNEIKYIASPDQKHFNVTDEALKDIDVALGDRIFSLEGTNNKET	360
QY	361	SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLFKEFFPEELKN	420
DB	361	SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLFKEFFPEELKN	420
QY	421	HGAYLGYTVTVSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSLSLTIHOAMRQOQIGSYF	480
DB	421	HGAYLGYTVTVSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSLSLTIHOAMRQOQIGSYF	480
QY	481	GSEITTSVDIDGQVTDVLLVAGAPMYNEGRERKGVYVYELQNRVFNCTLKDSSYQNA	540
DB	481	GSEITTSVDIDGQVTDVLLVAGAPMYNEGRERKGVYVYELQNRVFNCTLKDSSYQNA	540
QY	541	RFSSSTASVADLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG	600
DB	541	RFSSSTASVADLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG	600
QY	601	LQYFGCSIHQDLNEDGLDIAVGALGNVILWSRPVQINASLHFEPSKINIPIRDCCK	660
DB	601	LQYFGCSIHQDLNEDGLDIAVGALGNVILWSRPVQINASLHFEPSKINIPIRDCCK	660
QY	661	RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV	720
DB	661	RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV	720
QY	721	LLSSGGELCERINFHVLDTADYVKPTVFVSYSLEDPDHGPMDDGWPPTLVSVVPFWNG	780
DB	721	LLSSGGELCERINFHVLDTADYVKPTVFVSYSLEDPDHGPMDDGWPPTLVSVVPFWNG	780
QY	781	CNEDEHCVPDLVLDARSDDLPTAMEYQCVLRKPAQDCSAYTSLSFDTTFVIEESTQRVAV	840
DB	781	CNEDEHCVPDLVLDARSDDLPTAMEYQCVLRKPAQDCSAYTSLSFDTTFVIEESTQRVAV	840
QY	841	EATLENRGNASTVLNLSQSANLQFASLIQKEDSDGSECVNEERLQKQCNVSYPPFF	900
DB	841	EATLENRGNASTVLNLSQSANLQFASLIQKEDSDGSECVNEERLQKQCNVSYPPFF	900
QY	901	RAKAKVAFRLDSEFFSKSI FLHHLIELELAAGSDNSNRDSTKEDNAPLRFHLKYEADVLFT	960
DB	901	RAKAKVAFRLDSEFFSKSI FLHHLIELELAAGSDNSNRDSTKEDNAPLRFHLKYEADVLFT	960
QY	961	RSSLSLHYEVKLNLSLRYVDGTPPPSCIFRQNLGLPPIHGMMKTIPIATRGNRLI	1020
DB	961	RSSLSLHYEVKLNLSLRYVDGTPPPSCIFRQNLGLPPIHGMMKTIPIATRGNRLI	1020

Db 961 RSSLSHYEVKLNSSLERYDYGPPFCIFRIQNLGLFFPIHGMWMTIPIATRSNRL 1020
 QY 1021 KLRDLTDEANTSCNMGNSTEYRPTVEEDLRAPOLNHNDSVSNINCIRLVPNOEI 1080
 Db 1021 KLRDLTDEANTSCNMGNSTEYRPTVEEDLRAPOLNHNDSVSNINCIRLVPNOEI 1080
 QY 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPFFIFREEDPSRQIEFISKQEDWQV 1140
 Db 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPFFIFREEDPSRQIEFISKQEDWQV 1140
 QY 1141 PIWIIIVGSTGLGLLALLVIALRLKLGPRGARRRREGLDPTPKVLE 1188
 Db 1141 PIWIIIVGSTGLGLLALLVIALRLKLGPRGARRRREGLDPTPKVLE 1188

RESULT 2
 ABG12949
 ID ABG12949 standard; protein; 1189 AA.

XX ABG12949;
 XX
 XX 13-FEB-2002 (first entry)
 XX
 XX Novel human diagnostic protein #12940.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 XX 31-MAR-2000; 2000US-00540217.
 PR
 PR 23-AUG-2000; 2000US-00643167.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI: 2001-639362/73.
 DR
 DR N-PSDB; AAS77136.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PT
 XX
 XX Claim 20; SEQ ID NO 43308; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1189 AA;
 Query Match 78.1%; Score 928; DB 4; Length 1189;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDLPRGLVAVWALSLSLWPGFTDTFNMTRPRVPIQSGRTAFPGYTVQOHDISGNKWLAVGA 60
 Db 1 MDLPRGLVAVWALSLSLWPGFTDTFNMTRPRVPIQSGRTAFPGYTVQOHDISGNKWLAVGA 60
 QY 61 PLETNGVOKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNRLGLSLATPKDNSFLA 120
 Db 61 PLETNGVOKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNRLGLSLATPKDNSFLA 120
 QY 121 CSPLWHECGSSYTTGMCSSRVNSNFRSKTVAPALQRCQTYNDIVIVLDGNSIYPWVE 180
 Db 121 CSPLWHECGSSYTTGMCSSRVNSNFRSKTVAPALQRCQTYNDIVIVLDGNSIYPWVE 180
 QY 181 VQHFLINILAKFYIGPQIOGVVQVGEDVHVEHFLNDRSVKDVVEAASHIEQSGTET 240
 Db 181 VQHFLINILAKFYIGPQIOGVVQVGEDVHVEHFLNDRSVKDVVEAASHIEQSGTET 240
 QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLEKVIQOSERNVTRYAVAVL 300
 Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLEKVIQOSERNVTRYAVAVL 300
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVDEAALKDIVDALGDRIFSLGTKNKT 360
 Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVDEAALKDIVDALGDRIFSLGTKNKT 360
 QY 361 SFGLEMSOTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420
 Db 361 SFGLEMSOTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420
 QY 421 HGAYLGYTVTSVSSROGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQMRGQOIGSYF 480
 Db 421 HGAYLGYTVTSVSSROGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQMRGQOIGSYF 480
 QY 481 GSEITSDVIDDGVTDVLLVGA PMYFNEGRERGVVYVELQRNFVYNGTLKDSHSYQNA 540
 Db 481 GSEITSDVIDDGVTDVLLVGA PMYFNEGRERGVVYVELQRNFVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIPIHGFSGSILKTPKORITASELATG 600
 Db 541 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIPIHGFSGSILKTPKORITASELATG 600
 QY 601 LQYFCGSIHQGLDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Db 601 LQYFCGSIHQGLDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGGRFTNRAV 720
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGGRFTNRAV 720
 QY 721 LLSSGQELCERINFHVLDTADYVVPVTSVEYSLEDPDHGPMLDGHPPTTLRVSVPPWNG 780
 Db 721 LLSSGQELCERINFHVLDTADYVVPVTSVEYSLEDPDHGPMLDGHPPTTLRVSVPPWNG 780
 QY 781 CNEDEHCVDPDLVLDARSDDLPTAMEYCORVLKRPQDCSAYTSLSPDTTTFIESTRQRAV 840
 Db 781 CNEDEHCVDPDLVLDARSDDLPTAMEYCORVLKRPQDCSAYTSLSPDTTTFIESTRQRAV 840
 QY 841 EATLENRGENAYSTVLNLSQSANLQFASLQKEDSDSGSIECVNEERLQKQCVNSVPPFF 900
 Db 841 EATLENRGENAYSTVLNLSQSANLQFASLQKEDSDSGSIECVNEERLQKQCVNSVPPFF 900
 QY 901 RAKAKVAFRLDSEFSKSTFLHLELELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
 Db 901 RAKAKVAFRLDSEFSKSTFLHLELELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
 QY 961 RSSLSHYEVKLNSSLERYDYGPPFCIFRIQNLGLFFPIHGMWMTIPIATRSNRL 1020

Db 961 RSSLSHYEVKLSRLRYDGIQPPFCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020
Qy 1021 KLRDLTDE 1029
Db 1021 KLRDLTDE 1029
RESULT 3
ID AAU14231 standard; protein; 1188 AA.
XX AAU14231;
XX AAU14231;
XX 24-OCT-2001 (first entry)
XX Human novel protein #102.
XX Human; novel protein; Antianemic; osteopathic; antiinflammatory;
XX immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
XX anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
XX antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
XX thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
XX Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
XX tissue regeneration; immune disorder.
XX Homo sapiens.
XX WO200155437-A2.
XX 02-AUG-2001.
XX 25-JAN-2001; 2001WO-US002623.
XX 25-JAN-2000; 2000US-00491404.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-451939/48.
XX N-PSDB; AAS22536.
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
XX nervous system disorders, and for regenerating bone and cartilage.
XX Example 4; Page 578-591; 894pp; English.
XX The invention relates to polynucleotides encoding novel human proteins or
XX their active domains. The polypeptides, polynucleotides and antibodies
XX raised against the polypeptides are used in a method of treatment of a
XX mammal and prevention of disorders caused by the aberrant protein
XX expression or activity. The polypeptides can be used as molecular weight
XX markers, food supplements, and in antibody production. The polypeptides
XX are used to identify compounds which bind to the polypeptides.
XX Polynucleotides of the invention are used as probes and primers, for
XX sequencing, for chromosome or gene mapping, in the production of
XX recombinant proteins, and in generating anti-sense DNA or RNA and in gene
XX therapy. Polypeptides of the invention can be used to target drugs to a
XX tumour, in assays to determine biological activity, to raise
XX antibodies/elicit an immune response, to determine quantitative protein
XX levels, as tissue markers, and to isolate receptors or ligands.
XX Polypeptides of the invention may also be useful in treating platelet
XX disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX ligament and/or nerve tissue, wound healing, treating burns, promoting
XX the proliferation, differentiation and survival of stem cells, as a
XX contraceptive treating osteoporosis and osteoarthritis, anaemia,
XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
XX versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
XX diseases, nervous system disorders, and infection. The present sequence
XX represents a protein of the invention

XX Sequence 1188 AA;
SQ Query Match 77.9%; Score 926; DB 4; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MDLPRGLVWAWALSILWPGFTDTEFNDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVVGA 60
Db 1 MDLPRGLVWAWALSILWPGFTDTEFNDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVVGA 60
Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKDNSFLA 120
Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKDNSFLA 120
Qy 121 CSPLWSHECGSSVYTTGMSRVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYFWVE 180
Db 121 CSPLWSHECGSSVYTTGMSRVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYFWVE 180
Qy 181 VQHFLLNLIKFFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGTT 240
Db 181 VQHFLLNLIKFFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGTT 240
Qy 241 RTAFGIEFARSEAPQKGRKGAKKVMITVTDGSHSDSPDLEKVIQOSERNVTRYAVL 300
Db 241 RTAFGIEFARSEAPQKGRKGAKKVMITVTDGSHSDSPDLEKVIQOSERNVTRYAVL 300
Qy 301 GYNRGINPETFLNEIKYIASPDDKHPNVTDEAALKDIDVALGDRIFSLEGTNKT 360
Db 301 GYNRGINPETFLNEIKYIASPDDKHPNVTDEAALKDIDVALGDRIFSLEGTNKT 360
Qy 361 SFGLEMSQTGFSSHVEDGVLGAVGAYDWNGAVLKETSAGKVIPLRESYLKFEPEELKN 420
Db 361 SFGLEMSQTGFSSHVEDGVLGAVGAYDWNGAVLKETSAGKVIPLRESYLKFEPEELKN 420
Qy 421 HGAYLGYVTLSVSSRQGRVYVAGAPRHNHTGVILFTMHNRESLTHQAMRQOQISYF 480
Db 421 HGAYLGYVTLSVSSRQGRVYVAGAPRHNHTGVILFTMHNRESLTHQAMRQOQISYF 480
Qy 481 GSEITSDIDGCVTDVLLVGAPEMYNEGRERKGVYVYELRQNRFYNGTLDKSHSYQNA 540
Db 481 GSEITSDIDGCVTDVLLVGAPEMYNEGRERKGVYVYELRQNRFYNGTLDKSHSYQNA 540
Qy 541 RFGSSIASVDELNODSVNDVWVAGAPLEDHAGAIYIFHGFGRSILKTPQRTASELATG 600
Db 541 RFGSSIASVDELNODSVNDVWVAGAPLEDHAGAIYIFHGFGRSILKTPQRTASELATG 600
Qy 601 LQYFGCSIHQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEFSKINIFHRDCK 660
Db 601 LQYFGCSIHQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEFSKINIFHRDCK 660
Qy 661 RSRDATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
Db 661 RSRDATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
Qy 721 LLSGGQELCERINFHVLDTADVVKPTFSVSESLDPPDHGPMLDGWPPTTLRVSPFWNG 780
Db 721 LLSGGQELCERINFHVLDTADVVKPTFSVSESLDPPDHGPMLDGWPPTTLRVSPFWNG 780
Qy 781 CNEDEHCVPLDLARSDLPTAMEYQVRVLRPAQDCSAYTILSFDTTFFIESTQRVAV 840
Db 781 CNEDEHCVPLDLARSDLPTAMEYQVRVLRPAQDCSAYTILSFDTTFFIESTQRVAV 840
Qy 841 EATLENRGNASTVLNLSQSANLOFASLIQKEDSDGSLTECVNEERRLOKQVCNVSYPFF 900
Db 841 EATLENRGNASTVLNLSQSANLOFASLIQKEDSDGSLTECVNEERRLOKQVCNVSYPFF 900
Qy 901 RAKAKVAFRLDSEFSKISIFLHLEIELAAGSDNSERDSTKEDNAPLRFHLKYEADVLFT 960
Db 901 RAKAKVAFRLDSEFSKISIFLHLEIELAAGSDNSERDSTKEDNAPLRFHLKYEADVLFT 960
Qy 961 RSSLSHYEVKLSRLRYDGIQPPFCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020

```

Db 961 RSSLSHVEKPNSSLEKDYDGGPPFCIFRQNLGLFPIHGMWMTITPIATRSNRL 1020
Qy 1021 KLRDLTDEANTSCNMGNSYETPTPVEEDLRAPQNHNSNDVVSINCINILVFNQEI 1080
Db 1021 KLRDLTDEANTSCNMGNSYETPTPVEEDLRAPQNHNSNDVVSINCINILVFNQEI 1080
Qy 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFHSPFIPREDDPSROI 1128
Db 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFHSPFIPREDDPSROI 1128

RESULT 4
AAB50085
ID AAB50085 standard; protein; 1188 AA.
XX
AC AAB50085;
XX
DT 19-MAR-2001 (first entry)
XX
DE Human A259.
XX
KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key
FT Domain 1..1141
FT Peptide /label= Extracellular_domain
FT Protein 1..122
FT Domain /label= Signal_peptide
FT Domain 23..1188
FT Domain /label= Mature_protein
FT Domain 39..74
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#1
FT Domain 115..157
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#2
FT Domain 164..345
FT Domain /label= I_domain
FT Domain 367..392
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#3
FT Domain 421..455
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#4
FT Domain 478..516
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#5
FT Domain 540..575
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#6
FT Domain 602..640
FT Domain /label= Integrin_alpha_subunit_repeat_domain_#7
FT Domain 1142..1164
FT Domain /label= Transmembrane_domain
FT Domain 1165..1188
FT Domain /label= Cytoplasmic_domain
XX
PN WO200073339-A1.
XX
PD 07-DEC-2000.
XX
PE 15-MAY-2000; 2000WO-US013262.
XX
PR 28-MAY-1999; 99US-00322790.
PR 27-APR-2000; 2000US-00561263.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Pan Y, Lora JM;
XX
DR KPI; 2001-041142/05.
DR N-FSDB; AAC91901, AAC91902.
XX
PT Nucleic acid encoding alpha-integrin subunits, useful for treatment and
diagnosis of fibrosis, e.g. of the liver.

```

XX Claim 8; Fig 1; 164pp; English.

PS The present sequence is human integrin alpha subunit, A259. A259 is
 CC homologous with the alpha and alpha0 integrin subunits and is
 CC overexpressed in fibrosis. A259 is implicated in regulation of
 CC proliferation, differentiation and/or function of many different cell
 CC types. Inhibitors of A259 activity are useful for the treatment of liver
 CC disease, particularly fibrosis, and also fibrosis in other organs
 CC (specifically lung and kidney). In addition, A259 can be used for
 CC treatment and prevention of cancer, osteoporosis, acute myeloid
 CC leukaemia, HIV infection, and rheumatoid arthritis

XX Sequence 1188 AA;

Query Match 77.9%; Score 926; DB 4; Length 1188;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 MDLPRGLVAVKALSLWPGFTDTFNMDTRKPRVLPGRTPAFPGYTVQOHDISGNKWLAVGA 60
Db 1 MDLPRGLVAVKALSLWPGFTDTFNMDTRKPRVLPGRTPAFPGYTVQOHDISGNKWLAVGA 60
Qy 61 PLETNQYKTDGVYKCPVHGNCTKLNLRVTLSNYSERKDNVRLGLSLATNPKDNSFLA 120
Db 61 PLETNQYKTDGVYKCPVHGNCTKLNLRVTLSNYSERKDNVRLGLSLATNPKDNSFLA 120
Qy 121 CSPLWSHECGSSYYTTGWCSSRVNSNRFESKTAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 121 CSPLWSHECGSSYYTTGWCSSRVNSNRFESKTAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Qy 181 VOHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGGTET 240
Db 181 VOHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGGTET 240
Qy 241 RTAGIEFARSEAFQKGRKGAKKVMIVITDGSHPDLEKVIQOQSEKDNVTRYAVL 300
Db 241 RTAGIEFARSEAFQKGRKGAKKVMIVITDGSHPDLEKVIQOQSEKDNVTRYAVL 300
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALDKDIDVALDGRIFSLEGTNKNET 360
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALDKDIDVALDGRIFSLEGTNKNET 360
Qy 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKTSAGKVIPLRESYLKEFPEELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKTSAGKVIPLRESYLKEFPEELKN 420
Qy 421 HGAYLGYTVTTSVSSRQGRVYVAGAPRPNHTGKVIPLFTMNNRSLTIHQAMRQQIGSYF 480
Db 421 HGAYLGYTVTTSVSSRQGRVYVAGAPRPNHTGKVIPLFTMNNRSLTIHQAMRQQIGSYF 480
Qy 481 GSEITSVDIDGDGVTDVLLVGAPMYFNEGERKGVYVYELRQNRFYNGTLKDSHSYQNA 540
Db 481 GSEITSVDIDGDGVTDVLLVGAPMYFNEGERKGVYVYELRQNRFYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 600
Qy 601 LQYFGCSIHQGLDNLNEDGLIDLAGVAGNVLWSRPVQINASLHFEPSKINIHRDCK 660
Db 601 LQYFGCSIHQGLDNLNEDGLIDLAGVAGNVLWSRPVQINASLHFEPSKINIHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRVTPRAHLDEGGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRVTPRAHLDEGGDRFTNRAV 720
Qy 721 LLSGQELCERINPHVLDTADYKVPVTFSEVSELEDPDHGPMDDGWPTLVSVPFWNG 780
Db 721 LLSGQELCERINPHVLDTADYKVPVTFSEVSELEDPDHGPMDDGWPTLVSVPFWNG 780
Qy 781 CNEDEHCVPLDVLDARSDDLPTAMEYQORVLRKPAQCASYTISFDFTVFIIESTRORVAV 840
Db 781 CNEDEHCVPLDVLDARSDDLPTAMEYQORVLRKPAQCASYTISFDFTVFIIESTRORVAV 840

```

Db 781 CNEDEHCVPLDLVLDARSPLTAMEYQCVLRKPAQDCSAYTSLSDTTVFIIESTRQVAV 840
 Qy 841 EATLENRGENAYSTVLNISQSANIQFASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF 900
 Db 841 EATLENRGENAYSTVLNISQSANIQFASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF 900
 Qy 901 RAKAKVAFRLDSERFSKIFLHLEIELAAGSDSNERSDTKEDNVAPLRFHLKYEADVLF 960
 Db 901 RAKAKVAFRLDSERFSKIFLHLEIELAAGSDSNERSDTKEDNVAPLRFHLKYEADVLF 960
 Qy 961 RSSLSHYEVKNSLERYDGIQGFSCIFRIQNLGLFPIHGMWMMKITIPIATRSNRL 1020
 Db 961 RSSLSHYEVKNSLERYDGIQGFSCIFRIQNLGLFPIHGMWMMKITIPIATRSNRL 1020
 Qy 1021 KLRDFTLDEANTSCNMGNSTEYRPTVEEDLRRAPOLNHSNDVVSINIRLVPNQEI 1080
 Db 1021 KLRDFTLDEANTSCNMGNSTEYRPTVEEDLRRAPOLNHSNDVVSINIRLVPNQEI 1080
 Qy 1081 NFHLGLNLRSLKALKYKSKMIMVNAALQRFHSPFIREDPSRQI 1128
 Db 1081 NFHLGLNLRSLKALKYKSKMIMVNAALQRFHSPFIREDPSRQI 1128

RESULT 5
 AAU10551
 ID AAU10551 standard; protein; 1188 AA.
 XX AC AAU10551;
 DT 14-FEB-2002 (first entry)
 XX Human A259 polypeptide.

Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein; liver disease; fibrosis; lung; kidney; bone associated disorder; blood; cartilage associated disorder; haematopoietic disorder; bone marrow; immune related disease; apoptotic disorder; neuronal tissue disease; neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic; nephrotropic; immunomodulator; anti-inflammatory; neuroprotective; antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological; antidiabetic; anticonvulsant; antiparkinsonian.

Homo sapiens.

Key Location/Qualifiers
 Domain 1..1141
 Peptide 1..22 /note= "Extracellular domain"
 Protein 23..1188 /note= "Signal peptide"
 Domain 37..90 /note= "Mature human A259"
 Domain /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"
 Domain 115..157 /note= "Integrin alpha repeat domain"
 Domain 164..345 /note= "Integrin alpha repeat domain"
 Domain 367..392 /note= "Integrin alpha repeat domain"
 Domain 421..472 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"
 Domain 476..532 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"
 Domain 538..593 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"
 Domain 600..654 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"
 Domain 1142..1164

FT Domain /note= "Transmembrane domain"
 FT 1165..1188
 XX /note= "Cytoplasmic domain"
 PN W0200181414-A2.
 XX 01-NOV-2001.
 XX 27-APR-2001; 2001WO-US013516.
 XX 27-APR-2000; 2000US-00561263.
 PR (MILL-) MILLENNIUM PHARM INC.
 XX Pan Y, Lora J;
 PI WPI; 2002-041397/05.
 DR N-PSDB; AAS16873.
 XX New A259 nucleic acids and polypeptides, which comprise integrin alpha subunit, useful for diagnosing, preventing or treating e.g. liver disease, kidney or lung fibrosis, cancers, blood disorders or immune related diseases.
 XX Claim 9; Fig 1; 168pp; English.
 PS The invention relates to human and murine A259 nucleic acid molecules which encode secreted proteins with homology to integrin alpha subunits, specifically to integrin alpha 10. The A259 polypeptide and nucleic acid are useful for treating liver disease or fibrosis, particularly kidney fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also useful for diagnosing, preventing or treating cartilage and bone associated disorders (such as bone cancer, achondroplasia, myeloma, fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and osteoporosis), bone marrow, blood and haematopoietic disorders (such as acute myeloid leukaemia, haemophilia, anaemia and thalassaemia), immune related diseases (such as HIV, viral infections, cancers, T cell autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g. asthma and psoriasis), apoptotic disorders (such as systemic lupus erythematosus and insulin-dependent diabetes mellitus), diseases of the neuronal tissues (such as epilepsy and muscular dystrophy) and neurodegenerative diseases (such as Parkinson's disease and Huntington's disease). This sequence represents the human A259 polypeptide

Sequence 1188 AA;

Query Match 77.9%; Score 926; DB 5; Length 1188;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDLPRLGVAWALSMPGFTDTFNMDTRKPRVIPGSRRTAFFGYTVQOQHDISGNKMLVGA 60
 Db 1 MDLPRLGVAWALSMPGFTDTFNMDTRKPRVIPGSRRTAFFGYTVQOQHDISGNKMLVGA 60
 Qy 61 PLETNGVQKTGDVYKCPVHGNCTKLNLRVTLSNVSEKDNVRLGLSLATPKNSFLA 120
 Db 61 PLETNGVQKTGDVYKCPVHGNCTKLNLRVTLSNVSEKDNVRLGLSLATPKNSFLA 120
 Qy 121 CSLPWSHECGSSYYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180
 Db 121 CSLPWSHECGSSYYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180
 Qy 181 VOHFLINILKKFVIGPQIQVGVQYQGVYVHEFHNDYRSKDVVEAASHTEQRCGTET 240
 Db 181 VOHFLINILKKFVIGPQIQVGVQYQGVYVHEFHNDYRSKDVVEAASHTEQRCGTET 240
 Qy 241 RTAFGEFARSEAFQGGKGAQKVMIVITDGESHDSPLERKVIQOESRDNVTRAVAVL 300
 Db 241 RTAFGEFARSEAFQGGKGAQKVMIVITDGESHDSPLERKVIQOESRDNVTRAVAVL 300
 Qy 301 GYNNRGINPEFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLEGTNKNET 360
 Db 301 GYNNRGINPEFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLEGTNKNET 360

QY 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIVDALGDRIFSLGTTNKNET 360
 DB 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIVDALGDRIFSLGTTNKNET 360
 QY 361 SFGLSEMSQTGFSSHHVDEGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 DB 361 SFGLSEMSQTGFSSHHVDEGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQOQIGSYF 480
 DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQOQIGSYF 480
 QY 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGVVYVELQRNFRVYVNGTLKDSHSYQNA 540
 DB 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGVVYVELQRNFRVYVNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQITASELATG 600
 DB 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQITASELATG 600
 QY 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 DB 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSRGDTCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
 DB 661 RSRGDTCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
 QY 721 LLSGQELCRINPHVLDADYKVPVTFVSEYSLDPPDHGPMDDGWPITLRVSVFPMNG 780
 DB 721 LLSGQELCRINPHVLDADYKVPVTFVSEYSLDPPDHGPMDDGWPITLRVSVFPMNG 780
 QY 781 CNEDEHCVPLDLDARSDLPTAMEYQCVLRKPAQDCSAYTSLFDTTIVFIESTQRVAV 840
 DB 781 CNEDEHCVPLDLDARSDLPTAMEYQCVLRKPAQDCSAYTSLFDTTIVFIESTQRVAV 840
 QY 841 EATLENRGENAYTVLNISANLQFASLIQKEDSDGSEICVNEERRLQKVCNVSYPFF 900
 DB 841 EATLENRGENAYTVLNISANLQFASLIQKEDSDGSEICVNEERRLQKVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911

RESULT 7
ADE09956

ID ADE09956 standard; protein; 1188 AA.

XX AC ADE09956;

XX AC ADE09956;

DT 29-JAN-2004 (first entry)

XX Novel protein-related contig polypeptide sequence #544.

DE novel gene; novel protein; tissue marker; molecular weight marker;
 XX chromosome marker; Genetic disorder; contig.

XX Unidentified.

XX WO2003054152-A2.

XX PD 03-JUL-2003.

XX PF 10-DEC-2002; 2002WO-US039555.

XX PR 10-DEC-2001; 2001US-0339739P.

XX PR 11-DEC-2001; 2001US-0339453P.

XX PR 14-MAR-2002; 2002US-036091P.

XX PR 14-MAR-2002; 2002US-0365384P.

XX PR 12-APR-2002; 2002US-0372381P.

XX PR 12-APR-2002; 2002US-0372615P.

XX PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RI, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
 XX WPI; 2003-569235/53.
 XX New polynucleotides, useful for expressing recombinant proteins for
 PT analysis, characterization or therapeutic use, or as markers for tissues
 PT in which the corresponding protein is preferentially expressed.
 XX Disclosure; SEQ ID NO 3022; 1177pp; English.
 XX The invention comprises the amino acid and coding sequences of novel
 CC proteins. The DNA and protein sequences of the invention are useful as:
 CC markers for tissues in which the corresponding protein is preferentially
 CC expressed; as molecular weight markers on gels; as chromosome markers or
 CC tags; to identify chromosomes or to map related gene positions; and to
 CC compare with endogenous DNA sequences in patients to identify potential
 CC genetic disorders. The present amino acid sequence was used in the
 CC exemplification of the invention.

XX Sequence 1188 AA;

Query Match 76.7%; Score 911; DB 7; Length 1188;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVVAWALSPLPGFTDTFNMDTRKPRVTPGRTAFPGYTVQOHDISGNKWLAVGA 60

DB 1 MDLPRGLVVAWALSPLPGFTDTFNMDTRKPRVTPGRTAFPGYTVQOHDISGNKWLAVGA 60

QY 61 PLETNGYKGTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKDNSFLA 120

DB 61 PLETNGYKGTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKDNSFLA 120

QY 121 CSPLWSHECCSSVYTTGMCNRVNSNFRSKTAPALORCOTYMDIVILVDGNSIYFWVE 180

DB 121 CSPLWSHECCSSVYTTGMCNRVNSNFRSKTAPALORCOTYMDIVILVDGNSIYFWVE 180

QY 181 VQHEFLINILKKFYIGPGQIQVGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGTET 240

DB 181 VQHEFLINILKKFYIGPGQIQVGVVQYGEDVVFHFLNDYRSVKDVVEAASHIEQRGTET 240

QY 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQOSERDNTVRYAVVL 300

DB 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQOSERDNTVRYAVVL 300

QY 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIVDALGDRIFSLGTTNKNET 360

DB 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIVDALGDRIFSLGTTNKNET 360

QY 361 SFGLSEMSQTGFSSHHVDEGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420

DB 361 SFGLSEMSQTGFSSHHVDEGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420

QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQOQIGSYF 480

DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQOQIGSYF 480

QY 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGVVYVELQRNFRVYVNGTLKDSHSYQNA 540

DB 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGVVYVELQRNFRVYVNGTLKDSHSYQNA 540

QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQITASELATG 600

DB 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQITASELATG 600

QY 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660

601 LQYFGCSIHQGLDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
661 RSGRATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720
661 RSGRATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720
721 LLSGGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
721 LLSGGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
781 CNEDEHCVPLVLDARSPLTAMEYQORVLRPAQDCSAYTILSFDTTTFIESTQRVAV 840
781 CNEDEHCVPLVLDARSPLTAMEYQORVLRPAQDCSAYTILSFDTTTFIESTQRVAV 840
841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERLRQKQVCNVSYPFF 900
841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERLRQKQVCNVSYPFF 900
901 RAKAKVAFRLD 911
901 RAKAKVAFRLD 911
RESULT 8
AAB25582
ID AAB25582 standard; protein; 1189 AA.
XX
AC AAB25582;
XX
XX 21-NOV-2000 (first entry)
XX
DE ITGA11 protein encoded by human secreted protein gene #7.
XX
XX Secreted protein; immunosuppressant; anti-inflammatory; antithrombotic;
XX antithrombotic; dermatological; antiproliferative; antitumor; antitumor;
XX anticancer; vulvar; antiviral; antibacterial; antifungal;
XX immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
XX multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
XX Crohn's disease; nephritis; hyperproliferative disorder;
XX cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
XX melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
XX
OS Homo sapiens.
XX
XX WO200029435-A1.
XX
XX 25-MAY-2000.
XX
XX 27-OCT-1999; 99WO-US025031.
XX
XX 28-OCT-1998; 98US-0105971P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
XX Greene JM;
XX
XX WPI; 2000-387742/33.
XX
XX N-PSDB; AAA80612.
XX
XX Isolated nucleic acid molecules encoding human secreted proteins are used
XX for the prevention, amelioration and treatment of autoimmune,
XX inflammatory, hyperproliferative and cardiovascular disorders, cancer,
XX wounds, and infectious diseases.
XX
XX Claim 1; Fig 19A-F; 803pp; English.
XX
XX The present invention relates to 12 secreted human proteins and the
XX nucleotide sequences encoding them. The polynucleotide sequences given in
XX AAA80606-A80823 encode the 12 secreted protein sequences given in
XX AAB25576-B25593. The human secreted proteins have various activities
XX dependent on the tissues in which they are expressed. Examples of the
XX activities of the proteins include: immunosuppressant; anti-inflammatory;

antiarthritic; antirheumatic; dermatological; antiproliferative;
XX antiarteriosclerotic; anticancer; vulvar; antiviral; antibacterial;
XX and antifungal activity. The proteins, polypeptides, agonists and
XX antagonists may be used to treat prevent and/or diagnose various disease,
XX disorders and conditions examples of which include: immune disorders e.g.
XX Addison's disease, rheumatoid arthritis, dermatitis, and multiple
XX sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
XX Crohn's disease and nephritis; hyperproliferative disorders such as
XX arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
XX proteins and polynucleotide sequences may also be used in wound healing
XX and the treatment of infectious diseases. The human secreted protein gene
XX #7 and protein sequences are represented in sequences AAA80612 and
XX AAB25582. Secreted protein gene #7 is located at position chromosome 15
XX q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
XX the secreted protein gene#7
XX
XX Sequence 1189 AA;
Query Match 76.7%; Score 911; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDLPRGLVAVWALSPLPGFTDTENMDTRKPRVPGSRTAFPGVTVQOHDISGNKWLVVGA 60
DB 1 MDLPRGLVAVWALSPLPGFTDTENMDTRKPRVPGSRTAFPGVTVQOHDISGNKWLVVGA 60
QY 61 PLENGYQKTGDVYKCPVHGNCTKLNGLRVYLSNVSERKDNMRLGLSLATNPKNDSFLA 120
DB 61 PLENGYQKTGDVYKCPVHGNCTKLNGLRVYLSNVSERKDNMRLGLSLATNPKNDSFLA 120
QY 121 CSPLWSHECCSSYYTTCMSRVNSNFSKTVAPALQRCOTYMDIVLVLDGNSIYFWVE 180
DB 121 CSPLWSHECCSSYYTTCMSRVNSNFSKTVAPALQRCOTYMDIVLVLDGNSIYFWVE 180
QY 181 VOHFLNLKLFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVEAASHIEQGGTET 240
DB 181 VOHFLNLKLFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVEAASHIEQGGTET 240
QY 241 RTAFGIFFARSEAFQKGRKGAKKVMIVITDGSNHSFPLEKVTQOERDNTVRYAVL 300
DB 241 RTAFGIFFARSEAFQKGRKGAKKVMIVITDGSNHSFPLEKVTQOERDNTVRYAVL 300
QY 301 GYNRRGINPETELNEIKYIASDDPKDHFNVNTEAALKDIVDALGDRIFSLGNTKNET 360
DB 301 GYNRRGINPETELNEIKYIASDDPKDHFNVNTEAALKDIVDALGDRIFSLGNTKNET 360
QY 361 SFGLEMSQTGFSSHVEDGVLGAVGYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 420
DB 361 SFGLEMSQTGFSSHVEDGVLGAVGYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 420
QY 421 HGAVILGYTVTSVSSRQGRVYVAGAPRHNHTGKVLFTMHNRSITHOAMRGOQIGSYF 480
DB 421 HGAVILGYTVTSVSSRQGRVYVAGAPRHNHTGKVLFTMHNRSITHOAMRGOQIGSYF 480
QY 481 GSEITSVDIDGVTDLVLLGAPMYFNEGREGRGVYVYELQNRFPVYNGTLKDSHSYQNA 540
DB 481 GSEITSVDIDGVTDLVLLGAPMYFNEGREGRGVYVYELQNRFPVYNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNQSNDVYVAGAPLEDNHAGAVIIFHGFSGILKTKPKQITASELATG 600
DB 541 RFGSSIASVRDLNQSNDVYVAGAPLEDNHAGAVIIFHGFSGILKTKPKQITASELATG 600
QY 601 LQYFGCSIHQGLDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
DB 601 LQYFGCSIHQGLDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
QY 661 RSGRATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720
DB 661 RSGRATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720
QY 721 LLSGGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
DB 721 LLSGGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780

Db 721 LLSSGQELCERINPHVLDADYKVPVTFVSVEYSLEDDPDHGMDDGWPITLTVSVFWMNG 780
 QY 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLTDFDTTVFIESTRQVAV 840
 Db 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLTDFDTTVFIESTRQVAV 840
 QY 841 EATLENGENAYSTVLNISOSANLPASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF 900
 Db 841 EATLENGENAYSTVLNISOSANLPASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 Db 901 RAKAKVAFRLD 911
 RESULT 9
 ABR58364
 ID ABR58364 standard; protein; 1189 AA.
 XX AC ABR58364;
 XX DT 07-JUL-2003 (first entry)
 XX DE Human NOV2a.
 XX KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.
 OS Homo sapiens.
 XX PN W02003029423-A2.
 XX PD 10-APR-2003.
 XX PF 02-OCT-2002; 2002W0-US031358.
 XX PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327342P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 15-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 12-APR-2002; 2002US-0371972P.
 PR 12-APR-2002; 2002US-0371980P.
 PR 17-APR-2002; 2002US-0373261P.
 PR 19-APR-2002; 2002US-0373805P.
 PR 23-APR-2002; 2002US-0374738P.
 PR 16-MAY-2002; 2002US-0381101P.
 PR 17-MAY-2002; 2002US-0381635P.
 PR 29-MAY-2002; 2002US-0383830P.
 PR 01-OCT-2002; 2002US-00262839.
 (CURA-) CURAGEN CORP.
 PA Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
 XX Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;
 PI Rothenberg ME, Shinkets RA, Smithson G, Spyttek KA, Taupier RJ;
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
 XX WPI; 2003-381625/36.

DR N-PSDB; ACC72076.
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
 PT dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 1; Page 105; 487pp; English.
 XX The present invention relates to novel human NOV proteins and their
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
 CC proteins are useful in manufacturing a medicament for treating a syndrome
 CC associated with a human disease. The NOV proteins and coding sequences
 CC may be used to diagnose, treat or prevent metabolic disorders such as
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune
 CC disorders, haematopoietic disorders and various dyslipidaemias
 XX Sequence 1189 AA;
 SQ
 Query Match 76.7%; Score 911; DB 6; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDLPRGLVVAWALSPLPGFTDTFNMDTRKPRVPGSTAFFGYTVQOHDISGNKWLAVGA 60
 Db 1 MDLPRGLVVAWALSPLPGFTDTFNMDTRKPRVPGSTAFFGYTVQOHDISGNKWLAVGA 60
 QY 61 PLENGYQKTDGVYKCPVHNGCTKMLGRVTLNVSERKDNMRLGLSLATNPKDMSFLA 120
 Db 61 PLENGYQKTDGVYKCPVHNGCTKMLGRVTLNVSERKDNMRLGLSLATNPKDMSFLA 120
 QY 121 CSPLWSHECGSSYYTTCMCRSRVNSNFRFSKTVAPALQRCQTYNDIVILVDGNSIYPWVE 180
 Db 121 CSPLWSHECGSSYYTTCMCRSRVNSNFRFSKTVAPALQRCQTYNDIVILVDGNSIYPWVE 180
 QY 181 VQHFLINILKFKYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAAASHIQRCGTET 240
 Db 181 VQHFLINILKFKYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAAASHIQRCGTET 240
 QY 241 RTAFGIEFARSEAFQKGRGAKKMWIVITDGHSDSPDLKVIQOESRDNTRYAVAVL 300
 Db 241 RTAFGIEFARSEAFQKGRGAKKMWIVITDGHSDSPDLKVIQOESRDNTRYAVAVL 300
 QY 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLEGTNNKET 360
 Db 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLEGTNNKET 360
 QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKEFPPELQN 420
 Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKEFPPELQN 420
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHOAMRQQQIGSYF 480
 Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHOAMRQQQIGSYF 480
 QY 481 GSEITSDVDIDGDGVTVLLVGAAPMYNEGRGKVVYBELQRNFVYNGTLKDSHSYQNA 540
 Db 481 GSEITSDVDIDGDGVTVLLVGAAPMYNEGRGKVVYBELQRNFVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNQDSYNDVVGAPEDNHAGAIYIFHGFGRSILKTPQRITASELATG 600
 Db 541 RFGSSIASVRDLNQDSYNDVVGAPEDNHAGAIYIFHGFGRSILKTPQRITASELATG 600
 QY 601 LQYFGCSIHQGLDNLNEDGLIDLAVGALNAVILWSPVQVINASLHFEPSKINIFHRDCK 660
 Db 601 LQYFGCSIHQGLDNLNEDGLIDLAVGALNAVILWSPVQVINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYPTRAHLDEGGDRNTNRAV 720
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYPTRAHLDEGGDRNTNRAV 720
 QY 721 LJSSGQELCERINPHVLDADYKVPVTFVSVEYSLEDDPDHGMDDGWPITLTVSVFWMNG 780

Db 721 LLSGGQELCERINFHVDADYVVKVPTFSVEYSLEDPDHPGMLDDGWPTTLEVSVPFWNG 780
QY 781 CNEDEHCVDPDLVLDARSDLPAMEYCORVLRKPAQDCSAYTILSFDTTFIIESTRQAV 840
Db 781 CNEDEHCVDPDLVLDARSDLPAMEYCORVLRKPAQDCSAYTILSFDTTFIIESTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRIQKQVCNVSYPPF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRIQKQVCNVSYPPF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 10

ADA27054
ID ADA27054 standard; protein; 1189 AA.

AC ADA27054;

XX 20-NOV-2003 (first entry)

DE Human novel secreted protein from cDNA HOHBY69 #1.

XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;
KW hemostatic; gene therapy; cancer; inflammation; immune disorder;
KW neurological disorder; blood clotting disorder; food additive;
KW preservative; human; secreted protein.

XX Homo sapiens.

XX US2003055231-A1.

XX 20-MAR-2003.

XX 29-OCT-2001; 2001US-00984130.

XX 28-OCT-1998; 98US-0105971P.

XX 27-OCT-1999; 99WO-US025031.

XX 19-APR-2000; 2000US-0198407P.

XX 30-OCT-2000; 2000US-0243792P.

XX 18-APR-2001; 2001US-00836353.

XX (NIJJ/) NI J.

XX (YOUN/) YOUNG P E.

XX (KENN/) KENNY J J.

XX (OLSE/) OLSEN H S.

XX (MOOR/) MOORE P A.

XX (WEIY/) WEI Y.

XX (GREE/) GREENE J M.

XX (RUBE/) RUBEN S M.

XX (LIUD/) LIU D.

XX (CROC/) CROCKER P R.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

PI Ruben SM, Liu D, Crocker PR;

XX WPI: 2003-567103/53.

DR N-PSDB; ADA27036.

XX New human secreted nucleic acid molecules and polypeptides, useful for
PT preventing, treating, or ameliorating a medical condition, such as
PT cancer, inflammation, immune disorders, neurological and blood clotting
PT disorders.

XX Claim 11; Fig 19; 454pp; English.

XX The invention relates to an isolated nucleic molecule that is at least
CC 95% identical to 18 human cDNA sequences representing 12 novel genes
CC encoding secreted proteins or a polynucleotide fragment of the cDNA
CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic
CC variant of the polynucleotide having a polynucleotide capable of
CC hybridising under conditions the polynucleotide, where the polynucleotide
CC does not hybridise under stringent conditions to a nucleic acid molecule
CC having a nucleotide sequence of only A or T residues. Also included are
CC recombinant vectors, host cells (for producing the polypeptide), the
CC secreted polypeptide (comprising a sequence that is at least 95%
CC identical to a polypeptide fragment, domain, epitope, full-length
CC protein, variant, allelic variant or species homologue), antibodies that
CC specifically bind to the polypeptides, diagnosing, treating, preventing
CC or ameliorating a medical condition by administering the polynucleotide
CC or the polypeptide, the gene corresponding to the cDNA sequence and
CC identifying an activity in a biological assay (by expressing the cDNA
CC sequence in a cell, isolating the supernatant, and detecting an activity
CC in a biological assay and identifying the protein in the supernatant
CC having the activity). The polypeptides, nucleic acids and antibodies are
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition, for preventing, treating, or ameliorating a
CC medical condition, such as cancer, inflammation and other immune
CC disorders, neurological and blood clotting disorders (many examples are
CC given in the specification). The nucleic acids are also useful for
CC chromosome identification, radiation hybrid mapping or long-range
CC restriction mapping. The polypeptides and antibodies are useful for
CC providing immunological probes for differential identification of the
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,
CC agonist or antagonist may also be used as a food additive or preservative
CC to increase or decrease storage capabilities, fat content or other
CC nutritional components. The present is a secreted protein of the
CC invention.

XX Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 6; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVIFGSRTPAFGYTVQOHDISGNKWLVA 60
Db 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVIFGSRTPAFGYTVQOHDISGNKWLVA 60
QY 61 PLETNQYOKTGDVYKCPVHGNCTKLNLRVTLSNVSERKDNMRGLSLANPKDNSFLA 120
Db 61 PLETNQYOKTGDVYKCPVHGNCTKLNLRVTLSNVSERKDNMRGLSLANPKDNSFLA 120
QY 121 CSPLWSHECGSSYYTTGMCNRVNSNFRFSKTVA PALQRCQTYMDIIVLDGNSNIYPWE 180
Db 121 CSPLWSHECGSSYYTTGMCNRVNSNFRFSKTVA PALQRCQTYMDIIVLDGNSNIYPWE 180
QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVEAASHIEORGGTET 240
Db 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVEAASHIEORGGTET 240
QY 241 RTAFGIEFARSEAFQGGKRGAKKVMIVITDGEHSDSPDLEKVIQOQSERDNNVRYAV 300
Db 241 RTAFGIEFARSEAFQGGKRGAKKVMIVITDGEHSDSPDLEKVIQOQSERDNNVRYAV 300
QY 301 GYNNERGINPETFLNEIKYIASDDPDKHFNNVTDEAALKDIVDGLDIFLEGTKNKNET 360
Db 301 GYNNERGINPETFLNEIKYIASDDPDKHFNNVTDEAALKDIVDGLDIFLEGTKNKNET 360
QY 361 SPGLEMSQTRGSSHVVVEDGVLLGAVGYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420
Db 361 SPGLEMSQTRGSSHVVVEDGVLLGAVGYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420
QY 421 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTGKVLFTMHNNRSLTIHQAMRGOQIGSYF 480
Db 421 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTGKVLFTMHNNRSLTIHQAMRGOQIGSYF 480
QY 481 GSEITSVDIDGGVTDVLLVGAAPMYFNEGRGKVVYVELQNRVYVNGTLKDSHSYQNA 540
Db 481 GSEITSVDIDGGVTDVLLVGAAPMYFNEGRGKVVYVELQNRVYVNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNQDSYNDVVVGVVGAFLPDNHNAGAIYIFHGRGSILTKPKQRTASELATG 600

Db 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGFSGILKTPKQRTASELATG 600
 Qy 601 LOYFGCSIHGOLDNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Db 601 LOYFGCSIHGOLDNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGDFNRAV 720
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGDFNRAV 720
 Qy 721 LSSGQELCERINPHVLDADYKVPVTFSEYSLDHPDHPMLDDGWPPTLRVSVFPWNG 780
 Db 721 LSSGQELCERINPHVLDADYKVPVTFSEYSLDHPDHPMLDDGWPPTLRVSVFPWNG 780
 Qy 781 CNEDEHCVPLDVLDAKSDPTAMEYCORVLKPAQCSAYTLSFDTVTFTIESTRORVAV 840
 Db 781 CNEDEHCVPLDVLDAKSDPTAMEYCORVLKPAQCSAYTLSFDTVTFTIESTRORVAV 840
 Qy 841 EATLENRGENAYSTVLNISOSANLOFASLIQKEDSDGSEICVNEERRLOKQVCNVSYPFF 900
 Db 841 EATLENRGENAYSTVLNISOSANLOFASLIQKEDSDGSEICVNEERRLOKQVCNVSYPFF 900
 Qy 901 RAKAKVAFRLD 911
 Db 901 RAKAKVAFRLD 911

RESULT 11

ADE63570

ID ADE63570 standard; protein; 1189 AA.

XX ADE63570;

Dt 29-JAN-2004 (first entry)

XX Human Protein Q9URKX5, SEQ ID NO 9514.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WC2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-033347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 7; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALS LWPFGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLAVGA 60
 Db 1 MDLPRGLVAVWALS LWPFGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLAVGA 60
 Qy 61 PLETNGYQKTGDYVKCPVIHGNCTKLNLRVTLSNVSERKDNRLGLSLATNPKDNSFLA 120
 Db 61 PLETNGYQKTGDYVKCPVIHGNCTKLNLRVTLSNVSERKDNRLGLSLATNPKDNSFLA 120
 Qy 121 CSPLSWSEHCSSYYTTGMCSCRVSNSNFRFSKTVAPALQRCOTYMDIVLVDGNSIYPWVE 180
 Db 121 CSPLSWSEHCSSYYTTGMCSCRVSNSNFRFSKTVAPALQRCOTYMDIVLVDGNSIYPWVE 180
 Qy 181 VQFLINILKFKYIGPQIQVGVVQYGEDVYVHEFHLNDRSVKDVVEAASHIEQRGGTET 240
 Db 181 VQFLINILKFKYIGPQIQVGVVQYGEDVYVHEFHLNDRSVKDVVEAASHIEQRGGTET 240
 Qy 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQOSERNVTRYAVAVL 300
 Db 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQOSERNVTRYAVAVL 300
 Qy 301 GYNRRGINPETFLNEIKYIASPDQDKHPFNVTDEAALKDIDVALGDRIFSLGTTNKNET 360
 Db 301 GYNRRGINPETFLNEIKYIASPDQDKHPFNVTDEAALKDIDVALGDRIFSLGTTNKNET 360
 Qy 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWAVLKETSAGKVIPLRESYLKEPPELKN 420
 Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWAVLKETSAGKVIPLRESYLKEPPELKN 420
 Qy 421 HGAYLGYTVTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRLSLTIHQAMRQQQIGSYF 480
 Db 421 HGAYLGYTVTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRLSLTIHQAMRQQQIGSYF 480
 Qy 481 GSITSDVDIDGDVTDVLLVGAPEYNEGERKGVYVYELRQNRFTVNGTLKDSHYQNA 540
 Db 481 GSITSDVDIDGDVTDVLLVGAPEYNEGERKGVYVYELRQNRFTVNGTLKDSHYQNA 540
 Qy 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGFSGILKTPKQRTASELATG 600
 Db 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGFSGILKTPKQRTASELATG 600
 Qy 601 LOYFGCSIHGOLDNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Db 601 LOYFGCSIHGOLDNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGDFNRAV 720

Db 661 RSGDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720

Qy 721 LLSGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHPGMLDDGWPTTLRVSVFPWNG 780

Db 721 LLSGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHPGMLDDGWPTTLRVSVFPWNG 780

Qy 781 CNEDEHCVPDLVLDARSDLPATAMEYQCVLRKPAQDCSAYTILSFDTTTFIESTQRVAV 840

Db 781 CNEDEHCVPDLVLDARSDLPATAMEYQCVLRKPAQDCSAYTILSFDTTTFIESTQRVAV 840

Qy 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900

Db 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900

Qy 901 RAKAKVAFRLD 911

Db 901 RAKAKVAFRLD 911

RESULT 12

AD86584

ID AD86584 standard; protein; 1189 AA.

AC ADE86584;

XX

DT 29-JAN-2004 (first entry)

XX

DE Novel human secreted protein #7.

XX

KW human; secreted protein; cancer; liver disorder; hepatitis;

KW neural disorder; Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN US2003129685-A1.

XX

PD 10-JUL-2003.

XX

PF 18-APR-2001; 2001US-00836353.

XX

PR 28-OCT-1998; 98US-0105971P.

PR 27-OCT-1999; 99WO-US025031.

PR 19-APR-2000; 2000US-0198407P.

XX

PA (NIJJ/) NI J.

PA (YOUN/) YOUNG P E.

PA (KENN/) KENNY J J.

PA (OLSE/) OLSEN H S.

PA (MOOR/) MOORE P A.

PA (WEIY/) WEI Y.

PA (GREE/) GREENE J M.

PA (RUBE/) RUBEN S M.

XX

PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

PI Ruben SM;

XX

WPI: 2004-020335/02.

DR N-PSDB; ADE86566.

XX

XX New nucleic acid molecule, useful for preparing a medicament for

PT preventing, treating or ameliorating a medical condition e.g. cancer,

PT liver disorders or neural disorders.

XX

PS Claim 11; SEQ ID NO 35; 380pp; English.

XX

CC The invention relates to an isolated nucleic acid sequence, or its

CC allelic variant, a fragment of the cDNA sequence, or its fragment,

CC domain, epitope or species homologue. The nucleic acid is useful for

CC preparing a medicament for preventing, treating or ameliorating a medical

CC condition e.g., cancer, liver disorders such as hepatitis or neural

CC disorders such as Alzheimer's disease. The present sequence represents

CC the amino acid sequence of a novel human secreted protein.

XX

SQ Sequence 1189 AA;

Query Match

Best Local Similarity 76.7%; Score 911; DB 8; Length 1189;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPGRGVVAAWALSLEWFGFTDTFNMTRKPRVTPGSRTPAFGVTYQQHDSIGNKWLVVGA 60

Db 1 MDLPGRGVVAAWALSLEWFGFTDTFNMTRKPRVTPGSRTPAFGVTYQQHDSIGNKWLVVGA 60

Qy 61 PLENGYQKTDVYKCVIHHGNCCKLNLGRVTLNSYSEKDNMRLGLSLATNPKDNSFLA 120

Db 61 PLENGYQKTDVYKCVIHHGNCCKLNLGRVTLNSYSEKDNMRLGLSLATNPKDNSFLA 120

Qy 121 CSPLWSHECGSSYYTTCMCSRVSNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180

Db 121 CSPLWSHECGSSYYTTCMCSRVSNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180

Qy 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVYVHFFHNDYRSVKDVAASHIEQGGTET 240

Db 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVYVHFFHNDYRSVKDVAASHIEQGGTET 240

Qy 241 RTAFGIETFAFSEAFQKGRKGAKVMITITDGSNDSPLDKVITQSERDNVTRYAVAVL 300

Db 241 RTAFGIETFAFSEAFQKGRKGAKVMITITDGSNDSPLDKVITQSERDNVTRYAVAVL 300

Qy 301 GYNNRGINPETFLNEIKYITASDDPKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360

Db 301 GYNNRGINPETFLNEIKYITASDDPKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360

Qy 361 SFGLMSQTFSSHWVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420

Db 361 SFGLMSQTFSSHWVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420

Qy 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTGKVLFTMHNNSLTTHQAMRGQOIGSYF 480

Db 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTGKVLFTMHNNSLTTHQAMRGQOIGSYF 480

Qy 481 GSEITSDVIDGDTVDLLVGAPMYFNEGRERGVVYVYELQNFVYNGTLKDSHSYQNA 540

Db 481 GSEITSDVIDGDTVDLLVGAPMYFNEGRERGVVYVYELQNFVYNGTLKDSHSYQNA 540

Qy 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSLKTPKQITASELATG 600

Db 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSLKTPKQITASELATG 600

Qy 601 LQYFCGSIHGOLDLNEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660

Db 601 LQYFCGSIHGOLDLNEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720

Qy 721 LLSGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHPGMLDDGWPTTLRVSVFPWNG 780

Db 721 LLSGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHPGMLDDGWPTTLRVSVFPWNG 780

Qy 781 CNEDEHCVPDLVLDARSDLPATAMEYQCVLRKPAQDCSAYTILSFDTTTFIESTQRVAV 840

Db 781 CNEDEHCVPDLVLDARSDLPATAMEYQCVLRKPAQDCSAYTILSFDTTTFIESTQRVAV 840

Qy 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900

Db 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900

Qy 901 RAKAKVAFRLD 911

Db 901 RAKAKVAFRLD 911

RESULT 13

AAB25590

ID XX AAB25590 standard; protein; 1034 AA.
 AC AAB25590;
 DT 21-NOV-2000 (first entry)
 DE Protein encoded by human secreted protein gene #7 clone HOBY69.
 KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
 XX
 OS Homo sapiens.
 XX
 XX W0200029435-A1.
 XX
 XX 25-MAY-2000.
 XX
 XX 27-OCT-1999; 99WO-US025031.
 XX
 XX 28-OCT-1998; 98US-0105971P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
 XX Greene JM;
 XX WPI; 2000-387742/33.
 XX
 XX Isolated nucleic acid molecules encoding human secreted proteins are used
 XX for the prevention, amelioration and treatment of autoimmune,
 XX inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 XX wounds, and infectious diseases.
 XX
 XX Claim 1; Page 678-682; 803pp; English.
 XX
 XX The present invention relates to 12 secreted human proteins and the
 XX nucleotide sequences encoding them. The polynucleotide sequences given in
 XX AAA80606-A80623 encode the 12 secreted protein sequences given in
 XX AAB25576-B25593. The human secreted proteins have various activities
 XX dependent on the tissues in which they are expressed. Examples of the
 XX activities of the proteins include: immunosuppressant; anti-inflammatory;
 XX antiarthritic; antirheumatic; dermatological; antiproliferative;
 XX antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;
 XX and antifungal activity. The proteins, polypeptides, agonists and
 XX antagonists may be used to treat prevent and/or diagnose various disease,
 XX disorders and conditions examples of which include: immune disorders e.g.
 XX Addison's disease, rheumatoid arthritis, dermatitis, and multiple
 XX sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
 XX Crohn's disease and nephritis; hyperproliferative disorders such as
 XX paraneoplasias and purpura; cardiovascular disorders e.g. coronary
 XX arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
 XX proteins and polynucleotide sequences may also be used in wound healing
 XX and the treatment of infectious diseases. The human secreted protein gene
 XX #7 and protein sequences are represented in sequences AAA80612 and
 XX AAB25582. Secreted protein gene #7 is located at position chromosome 15
 XX q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
 XX the secreted protein gene#7
 XX
 XX Sequence 1034 AA;
 XX
 XX Query Match 68.2%; Score 810; DB 3; Length 1034;
 XX Best Local Similarity 99.9%; Pred. No. 0;
 XX Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDLPRGLVAVAWLSLWPGFTDTFNMDTRKPRVPGSRATFFGYTQQHDSGNKWLWVGA 60
 DB 1 MDLPRGLVAVAWLSLWPGFTDTFNMDTRKPRVPGSRATFFGYTQQHDSGNKWLWVGA 60

QY 61 PLENGYKTDGVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLGLSLATNPKDNSFLA 120
 DB 61 PLENGYKTDGVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLGLSLATNPKDNSFLA 120
 QY 121 CSPWASHCEGSSYYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVITVLGSSNIIYPWE 180
 DB 121 CSPWASHCEGSSYYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVITVLGSSNIIYPWE 180
 QY 191 VOHFLNLIKFFYIGPGQIOGVVQYGEDVVEHFLNDYVSVDVVEAAASHIQRGGTET 240
 DB 191 VOHFLNLIKFFYIGPGQIOGVVQYGEDVVEHFLNDYVSVDVVEAAASHIQRGGTET 240
 QY 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGSHSDSPDLKVIQCSERDNVTRYAVAVL 300
 DB 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGSHSDSPDLKVIQCSERDNVTRYAVAVL 300
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRFSLSGTNKNET 360
 DB 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRFSLSGTNKNET 360
 QY 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKFFPEELKN 420
 DB 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKFFPEELKN 420
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIIFTMNNRSLTIHQAMRQQIGSYF 480
 DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIIFTMNNRSLTIHQAMRQQIGSYF 480
 QY 481 GSEITSDVIDDGDVTDVLLVAGAPMYFNEGRERKGVVYELRQNFVYNGTLKDSHSYQNA 540
 DB 481 GSEITSDVIDDGDVTDVLLVAGAPMYFNEGRERKGVVYELRQNFVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKORITASELATG 600
 DB 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKORITASELATG 600
 QY 601 LOYFGCSIHGQDLNEDGLDLAGVAGNAVILWSRPVQINASLHFPESKINIHRDCK 660
 DB 601 LOYFGCSIHGQDLNEDGLDLAGVAGNAVILWSRPVQINASLHFPESKINIHRDCK 660
 QY 661 RSGRDATAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRVAV 720
 DB 661 RSGRDATAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRVAV 720
 QY 721 LLSGQELCERINPHVLDADYKVPVTFVSEYSELEDDHGMDDGWPTTLRVSVFVWNG 780
 DB 721 LLSGQELCERINPHVLDADYKVPVTFVSEYSELEDDHGMDDGWPTTLRVSVFVWNG 780
 QY 781 CNEDEHCVDPDLVDARSDLPTAMEYQVRLRKAQDCSAYTISFDTTVFIIESTORVAV 840
 DB 781 CNEDEHCVDPDLVDARSDLPTAMEYQVRLRKAQDCSAYTISFDTTVFIIESTORVAV 840
 QY 841 EATLENGENAYSTVNISQANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPFF 900
 DB 841 EATLENGENAYSTVNISQANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911
 RESULT 14
 ADA27062
 ID ADA27062 standard; protein; 1034 AA.
 XX
 XX ADA27062;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Human novel secreted protein from cDNA HOBY69 #2.
 XX
 XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;
 KW

KW hemostatic; gene therapy; cancer; inflammation; immune disorder;
KW neurological disorder; blood clotting disorder; food additive;
KW preservative; human; secreted protein.
XX Homo sapiens.
OS US2003055231-A1.
XX 20-MAR-2003.
XX 29-OCT-2001; 2001US-00984130.
XX 28-OCT-1998; 98US-0105971P.
XX 27-OCT-1999; 99WO-05025031.
XX 19-APR-2000; 2000US-0198407P.
XX 30-OCT-2000; 2000US-0243792P.
XX 18-APR-2001; 2001US-00836353.
XX (NIJ)/ NI J.
XX (YOUNG)/ YOUNG P E.
XX (KENN)/ KENNY J J.
XX (OLSE)/ OLSEN H S.
XX (MOOR)/ MOORE P A.
XX (WEIY)/ WEI Y.
XX (GREE)/ GREENE J M.
XX (RUBE)/ RUBEN S M.
XX (LIUD)/ LIU D.
XX (CROC)/ CROCKER P R.
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JW;
PI Ruben SM, Liu D, Crocker PR;
XX WPI; 2003-567103/53.
XX N-PSDB; ADA27044.
XX New human secreted nucleic acid molecules and polypeptides, useful for
PT preventing, treating, or ameliorating a medical condition, such as
PT cancer, inflammation, immune disorders, neurological and blood clotting
PT disorders.
XX Claim 11; Page 302-305; 454pp; English.
XX The invention relates to an isolated nucleic molecule that is at least
CC 9% identical to 18 human cDNA sequences representing 12 novel genes
CC encoding secreted proteins or a polynucleotide fragment of the cDNA
CC sequence contained in American Type Culture Collection (ATCC) deposit No.
CC defined in the specification, its species homologue, a variant or allelic
CC variant of the polynucleotide having a polynucleotide capable of
CC hybridising under conditions the polynucleotide, where the polynucleotide
CC does not hybridise under stringent conditions to a nucleic acid molecule
CC having a nucleotide sequence of only A or T residues. Also included are
CC recombinant vectors, host cells (for producing the polypeptide), the
CC secreted polypeptide (comprising a sequence that is at least 95%
CC identical to a polypeptide fragment, domain, epitope, full-length
CC protein, variant, allelic variant or species homologue), antibodies that
CC specifically bind to the polypeptides, diagnosing, treating, preventing
CC or ameliorating a medical condition by administering the polynucleotide
CC or the polypeptide, the gene corresponding to the cDNA sequence and
CC identifying an activity in a biological assay (by expressing the cDNA
CC sequence in a cell, isolating the supernatant, and detecting an activity
CC in a biological assay and identifying the protein in the supernatant
CC having the activity). The polypeptides, nucleic acids and antibodies are
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition, for preventing, treating, or ameliorating a
CC medical condition, such as cancer, inflammation and other immune
CC disorders, neurological and blood clotting disorders (many examples are
CC given in the specification). The nucleic acids are also useful for
CC chromosome identification, radiation hybrid mapping or long-range
CC restriction mapping. The polypeptides and antibodies are useful for
CC providing immunological probes for differential identification of the
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,
CC agonist or antagonist may also be used as a food additive or preservative
CC to increase or decrease storage capabilities, fat content or other

CC	nutritional components. The present is a secreted protein of the									
CC	invention.									
XX	SQ Sequence 1034 AA;									
QY	1	MDLPRGLVAVAWALSLWPGFTDTFNM	68.2%;	Score 810;	DB 6;	Length 1034;				
DB	1	MDLPRGLVAVAWALSLWPGFTDTFNM	Best Local Similarity	99.9%;	Pred. No. 0;					
QY	61	PLETNGYKQTDVYKCVIHNCTKLNGL	Matches 910;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps
DB	61	PLETNGYKQTDVYKCVIHNCTKLNGL								
QY	121	CSPLWSHECGSSYVTTGMC								
DB	121	CSPLWSHECGSSYVTTGMC								
QY	181	VOHFLINILKFFYIGPGQIQGVVQY								
DB	181	VOHFLINILKFFYIGPGQIQGVVQY								
QY	241	RTAFGIHFARSEAFQKGRGAKKVMIV								
DB	241	RTAFGIHFARSEAFQKGRGAKKVMIV								
QY	301	GYNRRGINPETFLNEIKYIASDPDD								
DB	301	GYNRRGINPETFLNEIKYIASDPDD								
QY	361	SFGLMSQTFSSHVVEDGVLGAVGYD								
DB	361	SFGLMSQTFSSHVVEDGVLGAVGYD								
QY	421	HGAYLGYVTWSVSSRQGRVYVAGAP								
DB	421	HGAYLGYVTWSVSSRQGRVYVAGAP								
QY	481	GSETSDVIDGDGVTDLVLLVGAPMY								
DB	481	GSETSDVIDGDGVTDLVLLVGAPMY								
QY	541	RFGSIAVRDLNODSYNDVVVVGAPLE								
DB	541	RFGSIAVRDLNODSYNDVVVVGAPLE								
QY	601	LQYFCGSIHGQDLNEDGLIDLAVGAL								
DB	601	LQYFCGSIHGQDLNEDGLIDLAVGAL								
QY	661	RSGRDATCLAAFLCFTPIFLAPHQTT								
DB	661	RSGRDATCLAAFLCFTPIFLAPHQTT								
QY	721	LLSSQQLCELRINFHVLDTADYVKP								
DB	721	LLSSQQLCELRINFHVLDTADYVKP								
QY	781	CNEDEHCVDPDLVDARSDLPTAMEY								
DB	781	CNEDEHCVDPDLVDARSDLPTAMEY								
QY	841	EATLENRGENAYSTVLNITSQANLQF								
DB	841	EATLENRGENAYSTVLNITSQANLQF								
QY	901	RAKAKVAFRLD 911								
DB	901	RAKAKVAFRLD 911								

RESULT 15
 ADE86592
 ID ADE86592 standard; protein; 1034 AA.
 AC ADE86592;
 DT 29-JAN-2004 (first entry)
 XX
 XX Novel human secreted protein #15.
 DE human; secreted protein; cancer; liver disorder; hepatitis;
 KW neural disorder; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 XX US2003129685-A1.
 XX
 XX 10-JUL-2003.
 XX
 XX 18-APR-2001; 2001US-00836353.
 XX
 XX 28-OCT-1998; 98US-0105971P.
 XX 27-OCT-1999; 99WO-US025031.
 XX 19-APR-2000; 2000US-0198407P.
 XX
 XX (NIJ/) NI J.
 XX (YOUN/) YOUNG P E.
 XX (KENN/) KENNY J J.
 XX (OLSE/) OLSEN H S.
 XX (MOOR/) MOORE P A.
 XX (WEI/) WEI Y.
 XX (GREE/) GREENE J M.
 XX (RUBE/) RUBEN S M.
 XX
 XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
 PI Ruben SM;
 XX
 XX WPI; 2004-020335/02.
 XX N-PSDB; ADE86574.
 XX
 XX New nucleic acid molecule, useful for preparing a medicament for
 PT preventing, treating or ameliorating a medical condition e.g. cancer,
 PT liver disorders or neural disorders.
 XX
 XX Claim 11; SEQ ID NO 43; 380pp; English.
 XX
 XX The invention relates to an isolated nucleic acid sequence, or its
 CC allelic variant, a fragment of the cDNA sequence, or its fragment,
 CC domain, epitope or species homologue. The nucleic acid is useful for
 CC preparing a medicament for preventing, treating or ameliorating a medical
 CC condition e.g., cancer, liver disorders such as hepatitis or neural
 CC disorders such as Alzheimer's disease. The present sequence represents
 CC the amino acid sequence of a novel human secreted protein.
 XX
 XX Sequence 1034 AA;
 SQ

Query Match
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRLVAVALSIPGFTDFNMTRKPRVTPGSRFAFGYVQOHDISGNKHLVYGA 60
 DB 1 MDLPRLVAVALSIPGFTDFNMTRKPRVTPGSRFAFGYVQOHDISGNKHLVYGA 60
 QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNGLRVLNSVSRKDNMRGLSLATNPKDNSFLA 120
 DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNGLRVLNSVSRKDNMRGLSLATNPKDNSFLA 120
 QY 121 CSPLWSHEGSSVYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180
 DB 121 CSPLWSHEGSSVYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180

QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDVFVAAASHIEQGGTET 240
 DB 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDVFVAAASHIEQGGTET 240
 QY 241 RTAFGIEFARSEAFQKGRKAKKVMIVITDGEHSDSPDLKVIQOQSERDNVRYAVAVL 300
 DB 241 RTAFGIEFARSEAFQKGRKAKKVMIVITDGEHSDSPDLKVIQOQSERDNVRYAVAVL 300
 QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLGKTKNET 360
 DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLGKTKNET 360
 QY 361 SFGLEMSQTGFSSHVHEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYKLPPEELKN 420
 DB 361 SFGLEMSQTGFSSHVHEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYKLPPEELKN 420
 QY 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRFNHTGKVLFTMHNRSLLTIHQAMRQQQIGSYF 480
 DB 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRFNHTGKVLFTMHNRSLLTIHQAMRQQQIGSYF 480
 QY 481 GSEITSDVIDGCVTDVLLVAGAPMYFNEGRERGVYVYELRQNRVYNGTLKDSHSYQNA 540
 DB 481 GSEITSDVIDGCVTDVLLVAGAPMYFNEGRERGVYVYELRQNRVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIPHGFRGSLTKTPKORITASELATG 600
 DB 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIPHGFRGSLTKTPKORITASELATG 600
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRVVQINASLHFEPSKINIFHRDCK 660
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRVVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATAFLAFLCFTPIFLAPHQITTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720
 DB 661 RSGRDATAFLAFLCFTPIFLAPHQITTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720
 QY 721 LLSSGQELCERINFHVLDTDADYVKPVTFSVEYSLEDPDHGPMDDGMPPTTLRVSVFPWNG 780
 DB 721 LLSSGQELCERINFHVLDTDADYVKPVTFSVEYSLEDPDHGPMDDGMPPTTLRVSVFPWNG 780
 QY 781 CNEDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQCSAYTLSFDITTVFIESTRQRAV 840
 DB 781 CNEDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQCSAYTLSFDITTVFIESTRQRAV 840
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF 900
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911

RESULT 16
 ABR58365
 ID ABR58365 standard; protein; 1120 AA.
 XX
 AC ABR58365;
 XX
 DT 07-JUL-2003 (first entry)
 XX
 XX Human NOV2b.
 XX
 KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antileptemic; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO2003029423-A2.

XX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031358.
XX
PR 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327342P.
PR 08-OCT-2001; 2001US-0327917P.
PR 08-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 28-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 12-APR-2002; 2002US-0371972P.
PR 12-APR-2002; 2002US-0371980P.
PR 17-APR-2002; 2002US-0373261P.
PR 19-APR-2002; 2002US-0373805P.
PR 23-APR-2002; 2002US-0374738P.
PR 16-MAY-2002; 2002US-0381101P.
PR 17-MAY-2002; 2002US-0381635P.
PR 29-MAY-2002; 2002US-0383830P.
PR 01-OCT-2002; 2002US-00262839.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E,
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W,
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK,
PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ,
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M,
XX
XX WPI; 2003-381625/36.
DR N-PSDB; ACC72077.
XX
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
PT dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 107; 487pp; English.
XX
XX The present invention relates to novel human NOV proteins and their
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
CC proteins are useful in manufacturing a medicament for treating a syndrome
CC associated with a human disease. The NOV proteins and coding sequences
CC may be used to diagnose, treat or prevent metabolic disorders such as
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
CC disorders such as Alzheimer's disease or Parkinson's disease, immune
CC disorders, haematopoietic disorders and various dyslipidaemias
XX
XX Sequence 1120 AA;
Query Match 63.4%; Score 753; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQTYMDIVILVLDGNSIPYVEVOHFLNLIKFFYIGPGQIQGVQYGEDVVHFEHLND 218
DB 90 CQTYMDIVILVLDGNSIPYVEVOHFLNLIKFFYIGPGQIQGVQYGEDVVHFEHLND 149
QY 219 YRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGEHSDSP 278
DB 150 YRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGEHSDSP 209
QY 279 DLEKVIQOESRDNTTRAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFNVTDEAAL 338
DB 210 DLEKVIQOESRDNTTRAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFNVTDEAAL 269

QY 339 KDIVDALGDRIFSLGNTKNKNETSGLEMSQTGFSHVHVDGVLLGAYDWMNGAVLKET 398
DB 270 KDIVDALGDRIFSLGNTKNKNETSGLEMSQTGFSHVHVDGVLLGAYDWMNGAVLKET 329
QY 399 SAGKVIPLRESYLKEPPEELKNHGAYLGYVTWTSVSSRQGRVYVAGAPRNFHTKVLFT 458
DB 330 SAGKVIPLRESYLKEPPEELKNHGAYLGYVTWTSVSSRQGRVYVAGAPRNFHTKVLFT 389
QY 459 MHNRSLLTIHOAMEGQOIGSYFGSEITSVDIDGQVTDVLLVGAPMYFNEGREGKVVY 518
DB 390 MHNRSLLTIHOAMEGQOIGSYFGSEITSVDIDGQVTDVLLVGAPMYFNEGREGKVVY 449
QY 519 ELRQNRFFVYNTLKDHSYQNAREFGSSIASYRDNLQDSYNDVVVGAPLEDNHAGAIYIFH 578
DB 450 ELRQNRFFVYNTLKDHSYQNAREFGSSIASYRDNLQDSYNDVVVGAPLEDNHAGAIYIFH 509
QY 579 GFRGSIILKTPKORITASELATGLOFGCSIHQGLDNLNEDGLIDLAVGALGNVILWSPV 638
DB 510 GFRGSIILKTPKORITASELATGLOFGCSIHQGLDNLNEDGLIDLAVGALGNVILWSPV 569
QY 639 VQINASHFEPFSKINIFHRDCKRGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 698
DB 570 VQINASHFEPFSKINIFHRDCKRGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 629
QY 699 ERYTPRAHLDEGGDRFTNRAVLLSSQGLCERINFHVDIADYVXKVPFVSFVSLDDPD 758
DB 630 ERYTPRAHLDEGGDRFTNRAVLLSSQGLCERINFHVDIADYVXKVPFVSFVSLDDPD 689
QY 759 HGPMLDDCGWPTTLRVSVFPFNGCNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCS 818
DB 690 HGPMLDDCGWPTTLRVSVFPFNGCNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCS 749
QY 819 AYTLSFDTTFFIESTQRVAEATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGS 878
DB 750 AYTLSFDTTFFIESTQRVAEATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGS 809
QY 879 IECVNEERRLOKOVNYSYPFFRAKAKVAFRLD 911
DB 810 IECVNEERRLOKOVNYSYPFFRAKAKVAFRLD 842
RESULT 17
AAU19663
ID AAU19663 standard; protein; 707 AA.
XX AC AAU19663;
XX DT 04-DEC-2001 (first entry)
XX DE Human novel extracellular matrix protein, Seq ID No 313.
XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
KW antianemic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antialzheimer's; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisenase; food additive.
XX OS Homo sapiens.
XX PN WO200155368-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001348.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214896P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226661P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231141P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-02311968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234744P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-025179P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465572/50.
N-PSDB; AAS31234.
Nucleic acid molecules encoding human secreted extracellular matrix proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers.

PS Claim 11; SEQ ID NO 313; 577pp; English.
XX The invention relates to isolated nucleic acid molecules encoding novel
CC human secreted extracellular matrix proteins (SPs). The polynucleotides
CC and proteins are used to prevent, treat a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For
CC example, disorders associated with decreased expression of SPs. The SP
CC polynucleotide or a vector expressing them may be administered to treat
CC diseases by gene therapy. Antisense molecules may be administered to down
CC regulate expression of SPs by binding with the cells own genes and
CC preventing their expression. The polynucleotides may also be used as DNA
CC probes in diagnostic assays. The SPs may also be used as antigens to
CC produce antibodies and to identify modulators (agonists and antagonists)
CC of the SPs. The anti-(sp) antibodies and antagonists may also be used to
CC down regulate expression and activity of SP and as diagnostic agents for
CC detecting the presence of SPs in samples. The disorders include for
CC example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency
CC virus) infections, anemia, rheumatoid arthritis and multiple sclerosis),
CC cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of
CC the breast or liver, Sezary syndrome and Gaucher's disease), neurological
CC diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-
CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC (e.g. corneal infections). Other uses include wound healing, maintenance
CC of organs before transplantation, support of cell culture of primary

Query Match 41.2%; Score 489; DB 4; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MDLPRGLVAVWALSINPFGFTDTFMDTRKPRVIGSRTAFFGYTVQOHDISGNKMLVGA 60
Db 17 MDLPRGLVAVWALSINPFGFTDTFMDTRKPRVIGSRTAFFGYTVQOHDISGNKMLVGA 76
Qy 61 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLNSVSRKDNRLGLSLATPNKNSFLA 120
Db 77 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLNSVSRKDNRLGLSLATPNKNSFLA 136
Qy 121 CSPLWSHECGSSYYTTCMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 137 CSPLWSHECGSSYYTTCMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
Qy 181 VQHLINILKFFIGPQIQGVVQYGEDVVFHFLNDVRSVDVVAASHIEQRGTTT 240
Db 197 VQHLINILKFFIGPQIQGVVQYGEDVVFHFLNDVRSVDVVAASHIEQRGTTT 256
Qy 241 RTAFGIEFARSAFQKGRGAKKVMIVITDGSHDSDPLEKVIQOSRDNVTRYAVAVL 300
Db 257 RTAFGIEFARSAFQKGRGAKKVMIVITDGSHDSDPLEKVIQOSRDNVTRYAVAVL 316
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGTTKNET 360
Db 317 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGTTKNET 376
Qy 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFFPEELKN 420
Db 377 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFFPEELKN 436
Qy 421 HGAYLGYTVTSVSSRGRRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
Db 437 HGAYLGYTVTSVSSRGRRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 496
Qy 481 GSEITSDVIDDGGVTDVLLVGA PMYFNEGRGKVVYELRQNFVYNGTLDKDSHYQNA 540
Db 497 GSEITSDVIDDGGVTDVLLVGA PMYFNEGRGKVVYELRQNFVYNGTLDKDSHYQNA 556
Qy 541 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGATVIFHGRGSLTKPKQRIITASELATG 600
Db 557 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGATVIFHGRGSLTKPKQRIITASELATG 616
Qy 601 LQYFGCSIHGOLDINEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIIFRDCK 660
Db 617 LQYFGCSIHGOLDINEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIIFRDCK 676

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707
RESULT 18
ABP47883
ID ABP47883 standard; protein; 707 AA.
AC ABP47883;
DT 23-AUG-2002 (first entry)
DE Human polypeptide SEQ ID NO 313.
KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antitickling; antianemic; antituberc; cancer;
KW antihemetic; hepatocytic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX Homo sapiens.
XX US2002042386-A1.
XX 11-APR-2002.
XX 17-JAN-2001; 2001US-00764870.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229333P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PT Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI: 2002-470713/50.
 DR N-PSDB; ABQ66558.
 XX
 PT New nucleic acid encoding human proteins, useful for diagnosis, treatment
 PT and prevention of e.g. osteoporosis, also related polypeptides and
 PT antibodies.
 XX
 PS Claim 11; SEQ ID NO 313; 235pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
 CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870
 XX
 SQ Sequence 707 AA;
 Query Match 41.2%; Score 489; DB 5; Length 707;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDLPRLVAVALSLLPFGFTDTFNMTRPRVTPGSRTPAFGVTYQCHDISGNKWLWGA 60
 DB 17 MDLPRLVAVALSLLPFGFTDTFNMTRPRVTPGSRTPAFGVTYQCHDISGNKWLWGA 76
 QY 61 PLETNGYQKTDGYKCPVHNGNCTKLNLRVLSNVSRKDNRLGLSLATNPKNSFLA 120
 DB 77 PLETNGYQKTDGYKCPVHNGNCTKLNLRVLSNVSRKDNRLGLSLATNPKNSFLA 136
 QY 121 CSPLWSEHCSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVLDGNSNLYPWVE 180
 DB 137 CSPLWSEHCSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVLDGNSNLYPWVE 196
 QY 181 VQHFNLINLKFFIGPQIQGVQVQYGEDVHVEHFLNDRYSVKDVEAASHIEQGGTET 240
 DB 197 VQHFNLINLKFFIGPQIQGVQVQYGEDVHVEHFLNDRYSVKDVEAASHIEQGGTET 256
 QY 241 RTAFGLFEARSEAFQGGKGAKKMIVITDGHSDSPLEKVIQOSERDNTVYAVAVL 300
 DB 257 RTAFGLFEARSEAFQGGKGAKKMIVITDGHSDSPLEKVIQOSERDNTVYAVAVL 316
 QY 301 GYNNRRGINPETFLNEIKYIASDPDDKHFNFVNTDEAALKDIDVALGDRIFSLEGTKNET 360
 DB 317 GYNNRRGINPETFLNEIKYIASDPDDKHFNFVNTDEAALKDIDVALGDRIFSLEGTKNET 376

QY 361 SFGLEMSOTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPEELKN 420
 DB 377 SFGLEMSOTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPEELKN 436
 QY 421 HGAYLGYTTSVSSRQGRVVVAGAPRNFHTKGVILFTMNNRSLTIHQAVRGQOIGSYF 480
 DB 437 HGAYLGYTTSVSSRQGRVVVAGAPRNFHTKGVILFTMNNRSLTIHQAVRGQOIGSYF 496
 QY 481 GSEITSVDIDGSDGYTDVLLVGPAPMYNEGRGKVVYVELRQNFVYNGTLKQSHSYQNA 540
 DB 497 GSEITSVDIDGSDGYTDVLLVGPAPMYNEGRGKVVYVELRQNFVYNGTLKQSHSYQNA 556
 QY 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIHFGRGSLTKPKQRTITASELATG 600
 DB 557 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIHFGRGSLTKPKQRTITASELATG 616
 QY 601 LOYFGCSIHGOLDLNEEDGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660
 DB 617 LOYFGCSIHGOLDLNEEDGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 676
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGI 691
 DB 677 RSGRDATCLAAFLCFTPIFLAPHFQTTVGI 707
 RESULT 19
 ADC10845
 ID ADC10845 standard; protein; 707 AA.
 XX AC ADC10845;
 XX DT 18-DEC-2003 (first entry)
 XX DE Human extracellular matrix protein from gene 38.
 XX KW Extracellular matrix protein; cytostatic; antibacterial; virucide;
 KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiac;
 KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
 KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
 KW neotropic; anti-allergic; cancer; bacterial infection; viral infection;
 KW neural disorder; immune system disorder; blood disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW inflammatory disorder; proliferative disorder; Human.
 XX OS Homo sapiens.
 XX FN US2003059875-A1.
 XX PD 27-MAR-2003.
 XX PF 19-APR-2002; 2002US-00125540.
 XX PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216847P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231444P.
PR 08-SEP-2000; 2000US-0231443P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235936P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 17-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764870.
XX
XX (HUNA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI
XX
XX WPI; 2003-743765/70.
XX N-PSDB; ADC10580.
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating, and/or preventing disorders, such as cancer, infections,
PT cardiovascular and inflammatory diseases.
XX
XX Claim 11; SEQ ID NO 313; 235pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA) encoding a human extracellular matrix protein, representing one of 161 novel genes. Also included are recombinant vectors, host cells (expressing the protein), the extracellular matrix proteins (including their fragments, epitopes and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or absence of a mutation in the nucleic acid and diagnosing a condition based on the presence or absence of the mutation), diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein), preventing, treating or ameliorating a medical condition by administering the nucleic acid or protein to a mammalian subject, identifying a binding partner to the

CC protein, the gene corresponding to the cDNA sequence, and identifying an
 CC activity in a biological assay (comprising expressing the nucleic acid in
 CC a cell, isolating the supernatant, detecting an activity in a biological
 CC assay and identifying the protein in the supernatant having the
 CC activity). The nucleic acids and proteins display the following
 CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,
 CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,
 CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 41.2%; Score 489; DB 7; Length 707;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MDLPRGLVWAWALSLPGFTDTFNMTRKPRVPGSRTAFGYTVQOHDISGNKWLWGA 60
 Db 17 MDLPRGLVWAWALSLPGFTDTFNMTRKPRVPGSRTAFGYTVQOHDISGNKWLWGA 76
 Qy 61 PLENTGQKTDGVVYKCPVHGNCTKLNGLRVTLNYSERKDNVRLGSLATPKNSFLA 120
 Db 77 PLENTGQKTDGVVYKCPVHGNCTKLNGLRVTLNYSERKDNVRLGSLATPKNSFLA 136
 Qy 121 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
 Db 137 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
 Qy 181 VQHFLLNLKXFFIGPQIQGVVQVGEDVVFHFLNDRSVKDVVEAASHIEQRGTTT 240
 Db 197 VQHFLLNLKXFFIGPQIQGVVQVGEDVVFHFLNDRSVKDVVEAASHIEQRGTTT 256
 Qy 241 RTAFGIEFARSEAFQKGRGAKKMWLITDGHSDSPLEKVIQOSRDNTRAVAVL 300
 Db 257 RTAFGIEFARSEAFQKGRGAKKMWLITDGHSDSPLEKVIQOSRDNTRAVAVL 316
 Qy 301 GYNRRGINPETFLNEIKYIASDPDDKHFNFVTDAAALKDIDALGDRIFSLGNTKNET 360
 Db 317 GYNRRGINPETFLNEIKYIASDPDDKHFNFVTDAAALKDIDALGDRIFSLGNTKNET 376
 Qy 361 SFGLMSQTFSSHVEDGVLLGAVAYDNGAVLAKETSAGKVIPLRESYLKEFFPEELKN 420
 Db 377 SFGLMSQTFSSHVEDGVLLGAVAYDNGAVLAKETSAGKVIPLRESYLKEFFPEELKN 436
 Qy 421 HGAYLGYTVTSVSSRQGRVYVACAPRNHTGKVLFTMNNRSLTIHQAMRGOQTGSYF 480
 Db 437 HGAYLGYTVTSVSSRQGRVYVACAPRNHTGKVLFTMNNRSLTIHQAMRGOQTGSYF 496
 Qy 481 GSITSDVDIDGCVTDVLLGAPWYNEGRERKVVYELRQNRVYNGTLKDSHYQNA 540
 Db 497 GSITSDVDIDGCVTDVLLGAPWYNEGRERKVVYELRQNRVYNGTLKDSHYQNA 556
 Qy 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGRFSILKTPKQRTIASL 600
 Db 557 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGRFSILKTPKQRTIASL 616
 Qy 601 LQYFGCSIHQDLNEDGLDIAVAGALGNVILWSPVQVQINASLHFFPSKINIFHRDCK 660
 Db 617 LQYFGCSIHQDLNEDGLDIAVAGALGNVILWSPVQVQINASLHFFPSKINIFHRDCK 676
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
 Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707

RESULT 20

AAU76854

ID AAU76854 standard; protein; 193 AA.

XX AAU76854;

XX AAU76854;

DT 21-MAY-2002 (first entry)

DE Human integrin alpha subunit Alpha 11 A domain.

XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;

KW Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;

KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
 KW ischaemia-reperfusion injury; immune complex; parasitic disease;
 KW antiinflammatory; vasotropic; antiparasitic; vulnerary; gene therapy.

OS Homo sapiens.

XX WO200209737-A1.

XX PD 07-FEB-2002.

XX PF 31-JUL-2001; 2001WO-US023957.

XX PR 31-JUL-2000; 2000US-0221950P.

XX PR 11-JAN-2001; 2001US-00758493.

XX PR 13-MAR-2001; 2001US-00805354.

XX PA (GEO) GEN HOSPITAL CORP.

XX PI Arnaout AM, Li R, Xiong J;

XX DR WPI; 2002-188687/24.

XX Novel high affinity integrin polypeptide useful for treating restenosis
 and parasitic diseases, comprises all or part of variant integrin alpha
 subunit A domain or variant integrin beta subunit A-like domain.

XX Example 2; Fig 5; 55pp; English.

XX The invention relates to a high affinity integrin polypeptide comprising
 all or part of a variant integrin alpha subunit A domain or a variant
 integrin beta subunit A-like domain. The polypeptide, preferably the
 CD11b alpha subunit A domain, where I at residue 332 has been replaced by
 G or A, F at residue 313 and A at residue 320 have been replaced by C, or
 V at residue 315 and A at residue 320 have been replaced by C, is useful
 for determining if a test compound is a candidate compound for binding to
 CD11b or for treating an inflammatory disorder, by contacting a test
 compound with the polypeptide and determining if the test compound binds
 to the polypeptide. The integrin subunits are useful for reducing
 skeletal muscle injury, for treating disorders caused by ischaemia-
 reperfusion injury, immune complexes, restenosis and parasitic diseases,
 to purify variant integrin polypeptide ligands and as bait proteins in
 two-hybrid or three-hybrid assays. This sequence represents the human
 integrin alpha subunit Alpha 11 A domain

XX Sequence 193 AA;

Query Match 16.2%; Score 193; DB 5; Length 193;
 Best Local Similarity 100.0%; Pred. No. 2.1e-182;

Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINLTKXFFIGPQIQGVVQVGEDVVFHFLND 218

Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINLTKXFFIGPQIQGVVQVGEDVVFHFLND 60

Qy 219 YRSVKDVVEAASHIEQRGTTTAFGIEFARSEAFQKGRGAKKMWLITDGHSDSP 278

Db 61 YRSVKDVVEAASHIEQRGTTTAFGIEFARSEAFQKGRGAKKMWLITDGHSDSP 120

Qy 279 DLEKVIQOSRDNTRAVAVLGYNNRRGINPETFLNEIKYIASDDDDKHFNFVTDAAAL 338

Db 121 DLEKVIQOSRDNTRAVAVLGYNNRRGINPETFLNEIKYIASDDDDKHFNFVTDAAAL 180

Qy 339 KDIVDALGDRIFS 351

Db 181 KDIVDALGDRIFS 193

RESULT 21

AAU76863

ID AAU76863 standard; protein; 193 AA.

XX AAU76863;

XX AAU76863;

DT 21-MAY-2002 (first entry)
DE Human integrin alpha subunit Alpha 11 variant A domain.
DE
XX
KW Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy;
KW mutuin.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 191
FT /note= "Wild-type Ile substituted by any other amino
FT acid"
XX
XX WO200209737-A1.
XX
XX 07-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US023957.
XX
XX 31-JUL-2000; 2000US-0221950P.
XX 11-JAN-2001; 2001US-00758493.
XX 13-MAR-2001; 2001US-00805354.
XX
XX (GEO) GEN HOSPITAL CORP.
XX
XX Arnaout AM, Li R, Xiong J;
XX
XX WPI; 2002-188687/24.
XX
XX Novel high affinity integrin polypeptide useful for treating restenosis
XX and parasitic diseases, comprises all or part of variant integrin alpha
XX subunit A domain or variant integrin beta subunit A-like domain.
XX
XX Claim 57; Page; 55pp; English.
XX
XX The invention relates to a high affinity integrin polypeptide comprising
XX all or part of a variant integrin alpha subunit A domain or a variant
XX integrin beta subunit A-like domain. The polypeptide, preferably the
XX CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX V at residue 315 and A at residue 320 have been replaced by C, is useful
XX for determining if a test compound is a candidate compound for binding to
XX CD11b or for treating an inflammatory disorder, by contacting a test
XX compound with the polypeptide and determining if the test compound binds
XX to the polypeptide. The integrin subunits are useful for reducing
XX skeletal muscle injury, for treating disorders caused by ischaemia-
XX reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX to purify variant integrin polypeptide ligands and as bait proteins in
XX two-hybrid or three-hybrid assays. This sequence represents a human
XX integrin alpha subunit Alpha 11 variant A domain. Note: This variant
XX sequence is not featured in the specification but has been derived from
XX the wild-type protein shown in AAU76854
XX
SQ Sequence 193 AA;
Query Match 16.0%; Score 190; DB 5; Length 193;
Best Local Similarity 100.0%; Pred. No. 2e-179;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 159 CQTYMDIVLDGNSIYPWVEVCHFLNLIKFKYIGPGQIQGVQVQYGEDVVEHFLND 218
Dy 1 CQTYMDIVLDGNSIYPWVEVCHFLNLIKFKYIGPGQIQGVQVQYGEDVVEHFLND 60
Qy 219 YRSKDVVEASHIEQKGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITDQESHDSP 278
Dy 61 YRSKDVVEASHIEQKGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITDQESHDSP 120
Qy 279 DLEKVIQOSERDNTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKDKHFFNVTDEAAL 338

Db 121 DLEKVIQOSERDNTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKDKHFFNVTDEAAL 180
Qy 339 KDIVDALGDR 348
Dy 181 KDIVDALGDR 190
RESULT 22
AAB50087
ID AAB50087 standard; protein; 1188 AA.
XX
XX AAB50087;
XX
XX 19-MAR-2001 (first entry)
XX Murine A259.
XX
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX
XX Mus sp.
XX
XX Location/Qualifiers
FH Key
FT Domain
FT /label= Extracellular_domain
FT Peptide
FT /label= Signal_peptide
FT Protein
FT /label= Mature_protein
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain
FT /label= I_domain
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain
FT /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain
FT /label= Transmembrane_domain
FT Domain
FT /label= Cytoplasmic_domain
XX
XX WO200073339-A1.
XX
XX 07-DEC-2000.
XX
XX 15-MAY-2000; 2000WO-US013262.
XX
XX 28-MAY-1999; 99US-00322790.
XX 27-APR-2000; 2000US-00561263.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX Pan Y, Lora JM;
XX
XX WPI; 2001-041142/05.
XX N-PSDB; AAC91904, AAC91905.
XX
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
XX diagnosis of fibrosis, e.g. of the liver.
XX
XX Claim 8; Fig 5; 164pp; English.
XX

CC The present sequence is murine integrin alpha subunit, A259. A259 is
 CC homologous with the alpha1 and alpha10 integrin subunits and is
 CC overexpressed in fibrosis. A259 is implicated in regulation of
 CC proliferation, differentiation and/or function of many different cell
 CC types. Inhibitors of A259 activity are useful for the treatment of liver
 CC disease, particularly fibrosis, and also fibrosis in other organs
 CC (specifically lung and kidney). In addition, A259 can be used for
 CC treatment and prevention of cancer, osteoporosis, acute myeloid
 CC leukaemia, HIV infection, and rheumatoid arthritis
 XX
 SQ Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GNCCTKLNGLGRVTLNVSERKDNMRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140
 DB 81 GNCCTKLNGLGRVTLNVSERKDNMRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140

QY 141 RVNSNFRFSKTVPALQRCQTYMDIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200
 DB 141 RVNSNFRFSKTVPALQRCQTYMDIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200

QY 201 VG 202
 DB 201 VG 202

RESULT 23
 AAU10552
 ID AAU10552 standard; protein; 1188 AA.
 AC AAU10552;
 DT 14-FEB-2002 (first entry)
 DE Murine A259 polypeptide.

XX Mouse; A259; integrin alpha subunit; integrin alpha 10; secreted protein;
 XX liver disease; fibrosis; lung; kidney; bone associated disorder; blood;
 XX cartilage associated disorder; haematopoietic disorder; bone marrow;
 XX immune related disease; apoptotic disorder; neuronal tissue disease;
 XX neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;
 XX nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;
 XX antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological;
 XX antidiabetic; anticonvulsant; antiparkinsonian.
 OS Mus musculus.

XX Key Location/Qualifiers
 FT Domain 1. .1141
 FT Peptide 1. .22 /note= "Extracellular domain"
 FT Protein 23. .1188 /note= "Signal peptide"
 FT Domain 39. .74 /note= "Mature murine A259"
 FT Domain 115. .157 /note= "Integrin alpha repeat domain"
 FT Domain 164. .345 /note= "Integrin alpha repeat domain"
 FT Domain 367. .392 /note= "I domain or Von Willebrand Factor type A domain"
 FT Domain 421. .455 /note= "Integrin alpha repeat domain"
 FT Domain 478. .516 /note= "Integrin alpha repeat domain"
 FT Domain 540. .575 /note= "Integrin alpha repeat domain"
 FT Domain 602. .640 /note= "Integrin alpha repeat domain"

FT Domain 1142. .1164 /note= "Transmembrane domain"
 FT Domain 1165. .1188 /note= "Cytoplasmic domain"
 XX WO200181414-A2.
 PD 01-NOV-2001.
 PF 27-APR-2001; 2001WO-US013516.
 PR 27-APR-2000; 2000US-00561263.
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Pan Y, Lora J;
 DR WPI; 2002-041397/05.
 DR N-PSDB; AAS16874.
 XX
 PT New A259 nucleic acids and polypeptides, which comprise integrin alpha
 PT subunit, useful for diagnosing, preventing or treating e.g. liver
 PT disease, kidney or lung fibrosis, cancers, blood disorders or immune
 PT related diseases.
 XX
 PS Claim 9; Fig 5; 168pp; English.
 XX
 CC The invention relates to human and murine A259 nucleic acid molecules
 CC which encode secreted proteins with homology to integrin alpha subunits,
 CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid
 CC are useful for treating liver disease or fibrosis, particularly kidney
 CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also
 CC useful for diagnosing, preventing or treating cartilage and bone
 CC associated disorders (such as bone cancer, achondroplasia, myeloma,
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and
 CC osteoporosis), bone marrow, blood and haematopoietic disorders (such as
 CC acute myeloid leukaemia, haemophilia, anaemia and thalassaemia), immune
 CC related diseases (such as HIV, viral infections, cancers, T cell
 CC autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.
 CC asthma and psoriasis), apoptotic disorders (such as systemic lupus
 CC erythematosus and insulin-dependent diabetes mellitus), diseases of the
 CC neuronal tissues (such as epilepsy and muscular dystrophy) and
 CC neurodegenerative diseases (such as Parkinson's disease and Huntington's
 CC disease). This sequence represents the murine A259 polypeptide
 XX
 SQ Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 5; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GNCCTKLNGLGRVTLNVSERKDNMRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140
 DB 81 GNCCTKLNGLGRVTLNVSERKDNMRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140

QY 141 RVNSNFRFSKTVPALQRCQTYMDIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200
 DB 141 RVNSNFRFSKTVPALQRCQTYMDIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200

QY 201 VG 202
 DB 201 VG 202

RESULT 24
 ADE08585
 ID ADE08585 standard; protein; 360 AA.
 AC ADE08585;
 XX
 DT 29-JAN-2004 (first entry)
 DE Novel protein (useful for identifying genetic disorders) #740.

XX novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder.
XX Unidentified.
XX WO2003054152-A2.
XX 03-JUL-2003.
XX 10-DEC-2002; 2002WO-US039555.
XX 10-DEC-2001; 2001US-0339739P.
XX 11-DEC-2001; 2001US-0339453P.
XX 14-MAR-2002; 2002US-0365091P.
XX 14-MAR-2002; 2002US-0365384P.
XX 12-APR-2002; 2002US-0372381P.
XX 12-APR-2002; 2002US-0372615P.
XX 22-APR-2002; 2002US-00128558.
XX 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI: 2003-569235/53.
XX N-PSDB; ADE07674.
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX Claim 20; SEQ ID NO 1651; 1177pp; English.
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
XX Sequence 360 AA;
SQ
Query Match 8.8%; Score 104; DB 7; Length 360;
Best Local Similarity 100.0%; Pred. No. 7.5e-94;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 251 SEAFQKGRGKAKVMIVITDGHSHSPDLEKVIQSSERDNTVRYAVLVGYNRRGINP 310
DB 232 SEAFQKGRGKAKVMIVITDGHSHSPDLEKVIQSSERDNTVRYAVLVGYNRRGINP 291
QY 311 ETFLENEIKYIASDPDDKHFFNVTDERALKDIVDALGRIFSLG 354
DB 292 ETFLENEIKYIASDPDDKHFFNVTDERALKDIVDALGRIFSLG 335
RESULT 25
ABG66673
ID ABG66673 standard; protein; 109 AA.
XX ABG66673;
XX 30-AUG-2002 (first entry)
XX Human novel polypeptide #8.
XX Human, inflammatory condition; shock; sepsis; immune response; cancer;
KW wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;

KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
XX fungal infection.
XX Homo sapiens.
XX OS
XX WO200244340-A2.
XX PN
XX 06-JUN-2002.
XX 30-NOV-2001; 2001WO-US047004.
XX PF
XX 30-NOV-2000; 2000US-00728952.
XX PR
XX (HYSE-) HYSEQ INC.
XX PA
XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
PI WPI: 2002-508509/54.
XX DR N-PSDB; ABK94897.
XX DR Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing.
XX PT
XX Claim 10; Page 572; 672pp; English.
XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention
XX Sequence 109 AA;
SQ
Query Match 7.5%; Score 89; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 2e-79;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDLPRGLVVAVALSILWPGFTDTFNNMDTRKPRVPGSRTAFFGYTVQQHDISGNKWLWGA 60
DB 1 MDLPRGLVVAVALSILWPGFTDTFNNMDTRKPRVPGSRTAFFGYTVQQHDISGNKWLWGA 60
QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNLG 89
DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNLG 89
RESULT 26
AAU19822
ID AAU19822 standard; protein; 103 AA.

XX AC AAU19822;
XX AC 06-DEC-2001 (first entry)
XX DT
XX DE Human novel extracellular matrix protein, Seq ID No 472.
XX KW
XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
XX KW anticancer; antirheumatic; antisclerotic; cardiac; vascular;
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
XX KW anti-alzheimers; immune/autoimmune disease; HIV infection; anaemia;
XX KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
XX KW cancers; hyperproliferative disease; breast neoplasm; melanoma;
XX KW Sezary syndrome; Gaucher's disease; neurological diseases;
XX KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
XX KW cardiac arrest; tachycardia; angina; infection; corneal infections;
XX KW wound healing; immunogen; gene therapy; antisense; food additive.
XX OS
XX KW Homo sapiens.
XX PN WO200155368-A1.
XX PD
XX DD 02-AUG-2001.
XX PF
XX PF 17-JAN-2001; 2001WO-US001348.
XX PR
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 11-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218390P.
XX PR 26-JUL-2000; 2000US-0220363P.
XX PR 28-JUL-2000; 2000US-0220364P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225475P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 14-AUG-2000; 2000US-0225759P.
XX PR 18-AUG-2000; 2000US-0226279P.
XX PR 22-AUG-2000; 2000US-0226681P.
XX PR 22-AUG-2000; 2000US-0226688P.
XX PR 22-AUG-2000; 2000US-0227182P.
XX PR 23-AUG-2000; 2000US-0227009P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 06-SEP-2000; 2000US-0229513P.
XX PR 06-SEP-2000; 2000US-0230437P.
XX PR 08-SEP-2000; 2000US-0230438P.
XX PR 08-SEP-2000; 2000US-0231242P.
XX PR 08-SEP-2000; 2000US-0231243P.
XX PR 08-SEP-2000; 2000US-0231244P.
XX PR 08-SEP-2000; 2000US-0231413P.
XX PR 08-SEP-2000; 2000US-0231414P.
XX PR 08-SEP-2000; 2000US-0232080P.
XX PR 08-SEP-2000; 2000US-0232081P.
XX PR 12-SEP-2000; 2000US-0231968P.
XX PR 14-SEP-2000; 2000US-0232397P.
XX PR 14-SEP-2000; 2000US-0232398P.
XX PR 14-SEP-2000; 2000US-0232399P.
XX PR 14-SEP-2000; 2000US-0232400P.
XX PR 14-SEP-2000; 2000US-0232401P.
XX PR 14-SEP-2000; 2000US-0233063P.
XX PR 14-SEP-2000; 2000US-0233064P.
XX PR 14-SEP-2000; 2000US-0233065P.
XX PR 21-SEP-2000; 2000US-0234223P.
XX PR 21-SEP-2000; 2000US-0234274P.
XX PR 25-SEP-2000; 2000US-0234997P.
XX PR 25-SEP-2000; 2000US-0234998P.
XX PR 26-SEP-2000; 2000US-0235484P.
XX PR 27-SEP-2000; 2000US-0235834P.
XX PR 27-SEP-2000; 2000US-0235836P.
XX PR 29-SEP-2000; 2000US-0236327P.
XX PR 29-SEP-2000; 2000US-0236367P.
XX PR 29-SEP-2000; 2000US-0236368P.
XX PR 29-SEP-2000; 2000US-0236369P.
XX PR 29-SEP-2000; 2000US-0236370P.
XX PR 02-OCT-2000; 2000US-0236802P.
XX PR 02-OCT-2000; 2000US-0237037P.
XX PR 02-OCT-2000; 2000US-0237038P.
XX PR 02-OCT-2000; 2000US-0237039P.
XX PR 13-OCT-2000; 2000US-0237040P.
XX PR 13-OCT-2000; 2000US-0239935P.
XX PR 13-OCT-2000; 2000US-0239937P.
XX PR 20-OCT-2000; 2000US-0240960P.
XX PR 20-OCT-2000; 2000US-0241221P.
XX PR 20-OCT-2000; 2000US-0241785P.
XX PR 20-OCT-2000; 2000US-0241787P.
XX PR 20-OCT-2000; 2000US-0241808P.
XX PR 20-OCT-2000; 2000US-0241809P.
XX PR 20-OCT-2000; 2000US-0241826P.
XX PR 01-NOV-2000; 2000US-0244617P.
XX PR 08-NOV-2000; 2000US-0246474P.
XX PR 08-NOV-2000; 2000US-0246475P.
XX PR 08-NOV-2000; 2000US-0246476P.
XX PR 08-NOV-2000; 2000US-0246477P.
XX PR 08-NOV-2000; 2000US-0246478P.
XX PR 08-NOV-2000; 2000US-0246523P.
XX PR 08-NOV-2000; 2000US-0246524P.
XX PR 08-NOV-2000; 2000US-0246525P.
XX PR 08-NOV-2000; 2000US-0246526P.
XX PR 08-NOV-2000; 2000US-0246527P.
XX PR 08-NOV-2000; 2000US-0246528P.
XX PR 08-NOV-2000; 2000US-0246609P.
XX PR 08-NOV-2000; 2000US-0246610P.
XX PR 08-NOV-2000; 2000US-0246611P.
XX PR 08-NOV-2000; 2000US-0246613P.
XX PR 17-NOV-2000; 2000US-0249207P.
XX PR 17-NOV-2000; 2000US-0249208P.
XX PR 17-NOV-2000; 2000US-0249209P.
XX PR 17-NOV-2000; 2000US-0249210P.
XX PR 17-NOV-2000; 2000US-0249211P.
XX PR 17-NOV-2000; 2000US-0249212P.
XX PR 17-NOV-2000; 2000US-0249213P.
XX PR 17-NOV-2000; 2000US-0249214P.
XX PR 17-NOV-2000; 2000US-0249215P.
XX PR 17-NOV-2000; 2000US-0249216P.
XX PR 17-NOV-2000; 2000US-0249217P.
XX PR 17-NOV-2000; 2000US-0249218P.
XX PR 17-NOV-2000; 2000US-0249244P.
XX PR 17-NOV-2000; 2000US-0249245P.
XX PR 17-NOV-2000; 2000US-0249264P.
XX PR 17-NOV-2000; 2000US-0249265P.


```

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233066P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-581633/65.
XX N-PSDB; ABK44005.
XX
XX New isolated nucleic acid encoding a protein for diagnosing, preventing,
XX treating or ameliorating medical conditions and used as food additives or
XX preservatives.
XX
XX Claim 9; SEQ ID NO 1193; 837pp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (III) encoded
XX by (I), are used to treat a medical conditions and in diagnosis of a
XX pathological condition. Disorders which are diagnosed or treated include
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease and
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX adenocarcinomas and irritable bowel syndrome reproductive system,
XX disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX leukaemia, disorders involving neovascularisation e.g. malignancies,
XX respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
XX acute kidney failure and blood related disorders e.g. myocardial
XX infarction. The polypeptides can also be used to aid wound healing and
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to
XX maintain organs before transplantation, for supporting cell culture of
XX primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities, fat content, lipid, protein,
Query Match 7.4%; Score 88; DB 4; Length 103;
Best Local Similarity 100.0%; Pred No. 1.9e-78;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVVHPEFLND 218
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVVHPEFLND 67
Qy 219 YRSVKDVEEAASHIEQRGGTETRTAFGI 246
Db 68 YRSVKDVEEAASHIEQRGGTETRTAFGI 95
RESULT 28
ABP48042
ID ABP48042 standard; protein; 103 AA.
XX ABP48042;
XX
XX 23-AUG-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 472.

```

XX Human; neutropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; anisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX Homo sapiens.
OS
XX
XX US2002042386-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-00764870.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214986P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218200P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2002-470713/50.
DR N-PSDB; ABQ66717.
XX New nucleic acid encoding human proteins, useful for diagnosis, treatment
PT and prevention of e.g. osteoporosis, also related polypeptides and
PT antibodies.
XX Claim 11; SEQ ID NO 472; 235pp + Sequence Listing; English.
PS
XX
XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870
XX
XX Sequence 103 AA;
SQ
Query Match 7.4%; Score 88; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.9e-78;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQTYMDIVILDGNSIYPWVEVOHFLINILKFFIYIGQIQGVGVQGVGVHFLND 218
D5 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 8 CQTYMDIVILDGNSIYPWVEVOHFLINILKFFIYIGQIQGVGVQGVGVHFLND 67
D5 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 219 YRSVKDVVEAASHIEQRGGTETRTAFGI 246
D5 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
D5 68 YRSVKDVVEAASHIEQRGGTETRTAFGI 95
RESULT 29
ADCI1004
ID ADCI1004 standard; protein; 103 AA.
XX
XX ADCI1004;
AC
XX 19-DEC-2003 (first entry)
DT
XX
XX Human protein from extracellular matrix gene 38 #2.
XX Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neutropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human.
XX Homo sapiens.
OS
XX
XX US2003059875-A1.
XX
XX 27-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR
XX 04-FEB-2000; 2000US-0180628P.
PR

PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217456P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0234989P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237039P.
PR 12-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246605P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764870.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-743765/70.

N-PSDB; ADC10739.

New isolated nucleic acids and polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.

XX PS Claim 11; SEQ ID NO 472; 235pp; English.

XX PS The invention relates to an isolated nucleic acid molecule (cDNA)

CC encoding a human extracellular matrix protein, representing one of 161

CC novel genes. Also included are recombinant vectors, host cells

CC (expressing the protein), the extracellular matrix proteins (including

CC their fragments, epitopes and homologues), an isolated antibody that

CC binds specifically to the protein, diagnosing a pathological condition or

CC susceptibility to a pathological condition (comprising determining the

CC presence or absence of a mutation in the nucleic acid and diagnosing a

CC condition based on the presence or absence of the mutation), diagnosing a

CC pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the

CC protein in a biological sample and diagnosing a condition based on the

CC presence or amount of expression of the protein), preventing, treating or

CC ameliorating a medical condition by administering the nucleic acid or

CC protein to a mammalian subject, identifying a binding partner to the

CC protein, the gene corresponding to the cDNA sequence, and identifying an

CC activity in a biological assay (comprising expressing the nucleic acid in

CC a cell, isolating the supernatant, detecting an activity in a biological

CC assay and identifying the protein in the supernatant having the

CC activity). The nucleic acids and proteins display the following

CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,

CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,

CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 7.4%; Score 88; DB 7; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.9e-78;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQYGEDVHPEFLND 218

Db 8 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQYGEDVHPEFLND 67

QY 219 YRSVKDVEAASHIEQRGGTETRTAFGI 246

Db 68 YRSVKDVEAASHIEQRGGTETRTAFGI 95

RESULT 30

ADB32059

ID ADB32059 standard; peptide; 42 AA.

AC ADB32059;

XX 04-DEC-2003 (first entry)

DE alphaA-integrin alpha subunit alphas1.

XX Integrin; alphav-beta3 integrin; RGD peptide ligand; modulator; agonist;

XX antagonist; alphas1.

XX Unidentified.

OS WO2003067219-A2.

PN 14-AUG-2003.

XX 07-FEB-2003; 2003WO-US003903.

XX 07-FEB-2002; 2002US-0354773P.

XX (GHO) GEN HOSPITAL CORP.

PA Arnaout AM;

XX WPI; 2003-663639/62.

XX Screening potential modulators of alphavbeta3 integrin useful to identify

PT agonists and antagonists uses computer model of three-dimensional

PT structure including a binding site and data from an alphavbeta3 integrin-

PT ligand complex.

XX PS Disclosure; Page 5c; 48pp; English.

XX PS The invention relates to a method for screening test compounds as

CC potential modulators of alphav-beta3 integrin using a computer model of

CC the three-dimensional structure of alphav-beta3 integrin which includes a

CC binding site. The model is based on atomic coordinates of defined alphav-

CC beta3 integrin amino acids obtained from the structure of a complex of

CC alphav-beta3 integrin with a known 'RGD peptide' ligand. The method is

CC useful to identify alphav-beta3 integrin ligands that, because they bind

CC to alphav-beta3 integrin, may be modulators e.g. agonists or antagonist

CC of alphav-beta3 integrin activity. It is useful to ascertain whether a

CC specific test compound is a potential modulator and especially to greatly

CC reduce numbers of compounds which must be further tested for their

CC ability to modulate alphav-beta3 integrin activity. The current sequence

CC represents the alphaA-integrin alpha subunit alphas1.

XX SQ Sequence 42 AA;

Query Match 3.5%; Score 42; DB 7; Length 42;

Best Local Similarity 100.0%; Pred. No. 3.7e-33;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 EAALKDIVDALGDRIFSLGINKNETSFGLEMSQTGFSSHV 376

Db 1 EAALKDIVDALGDRIFSLGINKNETSFGLEMSQTGFSSHV 42

RESULT 31

ABB72288

ID ABB72288 standard; protein; 545 AA.

XX ABB72288;

AC 04-APR-2002 (first entry)

DT Murine protein isolated from skin cells SEQ ID NO: 500.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;

XX developmental defect; inflammatory disease; dermatological; vulinary;

XX immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Mus sp.

OS WO200190357-A1.

PN 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

PI Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides

PT isolated from skin cells, useful for treating skin wounds, cancers,

PT growth and developmental defects, inflammatory diseases, or for

PT modulating immune responses.

XX Claim 4; Page 305-306; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs

CC isolated from human, murine and rat skin cell libraries. The sequences

CC can be used in the development of therapeutic agents useful in the

CC treatment of skin diseases, including skin wounds, cancer, growth

CC defects, developmental defects and inflammatory diseases. The proteins

CC have important roles in the induction of hair growth, cell proliferation

CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC polypeptide of the invention
XX
SQ Sequence 545 AA;

Query Match 3.5%; Score 42; DB 5; Length 545;
Best Local Similarity 100.0%; Pred. No. 3.5e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 123 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 164

RESULT 32
ABB72300
ID ABB72300 standard; protein; 688 AA.

XX ABB72300;
XX
XX
DT 04-APR-2002 (first entry)

XX Rat protein isolated from skin cells SEQ ID NO: 624.

DE Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnerary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Rattus sp.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ0000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

XX Kumble KD;

XX WPI; 2002-122020/16.

XX N-PSDB; ABL34985.

XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.

XX Claim 4; Page 388-390; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC polypeptide of the invention

XX Sequence 688 AA;

Query Match 3.5%; Score 42; DB 5; Length 688;
Best Local Similarity 100.0%; Pred. No. 4.3e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||

Db 266 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307

RESULT 33

ABB72289

ID ABB72289 standard; protein; 696 AA.

XX ABB72289;

XX 04-APR-2002 (first entry)

XX Rat protein isolated from skin cells SEQ ID NO: 501.

DE Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnerary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Rattus sp.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ0000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

XX Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.

XX Claim 4; Page 306-307; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC polypeptide of the invention

XX Sequence 696 AA;

Query Match 3.5%; Score 42; DB 5; Length 696;

Best Local Similarity 100.0%; Pred. No. 4.3e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807

|||||

Db 266 GWPTTLRVSVFFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307

RESULT 34

AAM39928

ID AAM39928 standard; protein; 757 AA.

XX AAM39928;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3073.

XX

XX	Novel human diagnostic protein #12941.
DE	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	food supplement; medical imaging; diagnostic; genetic disorder.
KW	
XX	Homo sapiens.
OS	
XX	WO2001175067-A2.
PN	
XX	11-OCT-2001.
PD	
XX	30-MAR-2001; 2001WO-US008631.
PF	
XX	31-MAR-2000; 2000US-00540217.
PR	
XX	23-AUG-2000; 2000US-00649167.
PR	
XX	(HYSE-) HYSEQ INC.
PA	
XX	Drmanac RT, Liu C, Tang YT;
PI	
XX	WPI; 2001-639362/73.
DR	
XX	N-PSDB; AAS71137.
DR	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
PT	
XX	Claim 20; SEQ ID NO 43309; 103pp; English.
PS	
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 117 AA;
SQ	

```
Query Match      2.98; Score 34; DB 4; Length 117;
Best Local Similarity 100.0%; Pred.No. 7.9e-25;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

ID	AAB25614 standard; peptide; 33 AA.
XX	AC AAB25614;
XX	AC AAB25614;
XX	DT 21-NOV-2000 (first entry)
XX	DE Human secreted protein ITGAL peptide encoded by secreted protein #6.

CC providing immunological probes for differential identification of the
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,
CC agonist or antagonist may also be used as a food additive or preservative
CC to increase or decrease storage capabilities, fat content or other
CC nutritional components. The present is a peptide motif from a secreted
CC protein of the invention.
XX
SQ Sequence 33 AA;

Query Match 2.8%; Score 33; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 TNGYQKTGDYKCPVHGNCTKLNLRVTLNV 96
DB 1 TNGYQKTGDYKCPVHGNCTKLNLRVTLNV 33

RESULT 38

AD86651
ID ADE86651 standard; peptide; 33 AA.

AC ADE86651;

XX 29-JAN-2004 (first entry)

DT Transmembrane domain.

DE human; secreted protein; cancer; liver disorder; hepatitis;
KW neural disorder; Alzheimer's disease; transmembrane domain.

XX Homo sapiens.

XX US2003129685-A1.

XX 10-JUL-2003.

XX 18-APR-2001; 2001US-00836353.

XX 28-OCT-1998; 98US-0105971P.

XX 27-OCT-1999; 99WO-US025031.

XX 19-APR-2000; 2000US-0198407P.

XX (NIJ/) NI J.

XX (YOUN/) YOUNG P E.

XX (KEND/) KENDY J J.

XX (OLSE/) OLSEN H S.

XX (MOOR/) MOORE P A.

XX (WEIY/) WEI Y.

XX (GREE/) GREENE J M.

XX (RUBE/) RUBEN S M.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

PI Ruben SM;

XX WPI; 2004-020335/02.

XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.

XX Disclosure; SEQ ID NO 102; 380pp; English.

XX The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of a transmembrane domain.

XX Sequence 33 AA;

Query Match 2.0%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 2.8%; Score 33; DB 8; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 TNGYQKTGDYKCPVHGNCTKLNLRVTLNV 96
DB 1 TNGYQKTGDYKCPVHGNCTKLNLRVTLNV 33

RESULT 39

AAB30927
ID AAB30927 standard; peptide; 24 AA.

XX AC AAB30927;

XX 02-APR-2001 (first entry)

DE Peptide derived from a human alpha1 integrin chain.

XX Human; integrin; alpha1 subunit; fibroblast; muscle cell; chondrocyte;
KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
KW osteoporosis; cartilage damage; bone damage; cartilage.

XX Homo sapiens.

XX WO200075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SE001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Claim 23; Page 40; 79pp; English.

XX The present sequence is derived from the cytoplasmic domain of the human
CC integrin subunit, designated alpha11. The alpha11 polynucleotide and
CC polypeptide are useful as markers of cell target molecules, such as
CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
CC derived cells or stem cells. They are also used for determining the
CC differential stage of cells during differentiation, development in
CC pathological conditions, in tissue regeneration, in transplantation or in
CC therapeutic and physiological repair of tissues. The pathological
CC conditions involving subunit alpha11 are selected from damage of cells,
CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
CC osteoarthritis and osteoporosis, damage of cartilage and bone, and
CC cartilage and bone diseases. The polypeptide is useful for detecting the
CC formation of cartilage during embryonic development, for detecting
CC physiological therapeutic repair of cartilage and muscle, for selection
CC and analysis, or for sorting, isolating or purification of chondrocytes
CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
CC during transplantation of cartilage or chondrocytes during
CC transplantation of cartilage or chondrocytes, respectively, or of muscle
CC or muscle cells during transplantation of muscle or muscle cells,
CC respectively, and for studies of differentiation or chondrocytes or
CC muscle cells

XX Sequence 24 AA;

Query Match 2.0%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1165 KLGFFRSARRRRPGLDPTPKVLE 1188
 Db 1 KLGFFRSARRRRPGLDPTPKVLE 24

RESULT 40

AAAY76112
 ID AAY76112 standard; protein; 158 AA.

XX AC AAY76112;

XX DT 27-MAR-2000 (first entry)

XX DE Rat integrin homologue, SEQ ID NO:391.

XX KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.

XX OS Rattus sp.

XX PN WO9955865-A1.

XX PD 04-NOV-1999.

XX PF 29-APR-1999; 99WO-NZ0000051.

XX PR 29-APR-1998; 98US-00069726.

XX PR 09-NOV-1998; 98US-00188930.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Strachan L, Sleeman M, Watson JD, Orrust R, Kumble A, Murison JG;
 XX DR N-PSDB; AAZ61820.

XX PT Novel polynucleotides useful for the treatment of various conditions
 XX including wounds and cancer.
 XX Claim 4; Page 221; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
 CC skin inflammation, to modulate epithelial cell growth and to inhibit
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat
 CC growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
 CC cDNA sequences derived from several mouse, rat or human skin cell types.
 CC Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
 CC are proteins with an N-terminal signal sequence, indicating that they are
 CC secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
 CC AAY76121-Y76122 are proteins with one or more putative transmembrane
 CC domains

XX SQ Sequence 158 AA;

Query Match 1.8%; Score 21; DB 3; Length 158;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 636 RPVVQINASLHFPSPKINIFH 656

Db 57 RPVVQINASLHFPSPKINIFH 77

RESULT 41

AAAB56051
 ID AAB56051 standard; protein; 158 AA.

XX AC AAB56051;

XX DT 08-MAR-2001 (first entry)

XX DE Skin cell protein, SEQ ID NO: 391.

XX KW Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
 KW neuroprotective; vulnery; immunomodulatory; vaccine;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.

XX OS Rattus sp.

XX PN WO200069884-A2.

XX PD 23-NOV-2000.

XX PF 15-MAY-2000; 2000WO-NZ0000075.

XX PR 14-MAY-1999; 99US-00312283.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD;
 XX PI Murison JG;

XX DR WPI; 2001-007495/01.

XX DR N-PSDB; AAC99753.

XX PT New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases.
 XX Claim 4; Page 296-297; 352pp; English.

XX The present sequence is a polypeptide which is expressed in mammalian
 CC skin cells. The polypeptide is useful for stimulating keratinocyte growth
 CC and motility, inhibiting the growth of cancer cells, modulating
 CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,
 CC modulating skin inflammation, stimulating the growth of epithelial cells,
 CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to
 CC leukocytes, and treating inflammatory disease, cancer and neurological
 CC diseases. The polynucleotide can be used as a marker, in the
 CC identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns

XX SQ Sequence 158 AA;

Query Match 1.8%; Score 21; DB 4; Length 158;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 636 RPVVQINASLHFPSPKINIFH 656

Db 57 RPVVQINASLHFPSPKINIFH 77

RESULT 42

ABB72251
 ID ABB72251 standard; protein; 158 AA.

XX AC ABB72251;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 391.

XX

KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnerary;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
 XX
 OS Rattus sp.
 XX
 FN WO200190357-A1.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-NZ000099.
 XX
 PR 24-MAY-2000; 2000US-0206650P.
 PR 25-JUL-2000; 2000US-0221232P.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
 PI Kumble KD;
 XX
 DR WPI; 2002-122020/16.
 XX
 PT New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses.
 XX
 PS Example 2; Page 246; 466pp; English.
 XX
 CC The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention
 XX
 SQ Sequence 158 AA;
 Query Match 1.8%; Score 21; DB 5; Length 158;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 RPVVQINASLHPFPSPKINIFH 656
 DB 57 RPVVQINASLHPFPSPKINIFH 77
 RESULT 43
 AAY76111
 ID AAY76111 standard; protein; 85 AA.
 XX
 AC AAY76111;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Rat integrin homologue, SEQ ID NO:390.
 XX
 KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnerary.
 XX
 OS Rattus sp.
 XX
 FN WO9955865-A1.
 XX
 PD 04-NOV-1999.
 XX
 PF 29-APR-1999; 99WO-NZ000051.
 XX
 PR 29-APR-1998; 98US-00069726.
 PR 09-NOV-1998; 98US-00188930.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 PI WPI; 2000-072177/06.
 DR N-PSDB; AAZ61819.
 XX
 PT Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer.
 XX
 PS Claim 4; Page 221; 235pp; English.
 XX
 CC The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
 CC skin inflammation, to modulate epithelial cell growth and to inhibit
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat
 CC growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
 CC cDNA sequences derived from several mouse, rat or human skin cell types.
 CC Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
 CC are proteins with an N-terminal signal sequence, indicating that they are
 CC secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
 CC AAY76121-Y76122 are proteins with one or more putative transmembrane
 CC domains
 XX
 SQ Sequence 85 AA;
 Query Match 1.4%; Score 17; DB 3; Length 85;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 528 NGTLKDSHSYQNAFSGS 544
 DB 28 NGTLKDSHSYQNAFSGS 44
 RESULT 44
 AAB56050
 ID AAB56050 standard; protein; 85 AA.
 XX
 AC AAB56050;
 XX
 DT 08-MAR-2001 (first entry)
 XX
 DE Skin cell protein, SEQ ID NO: 390.
 XX
 KW Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; neurotropic;
 KW neuroprotective; vulnerary; immunomodulatory; vaccine; inhibition;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.
 XX
 OS Rattus sp.
 XX
 FN WO200059884-A2.
 PD 23-NOV-2000.
 XX
 PF 15-MAY-2000; 2000WO-NZ000075.
 XX
 PR 14-MAY-1999; 99US-00312283.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX

PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;
PI Murison JG;
XX
XX
XX WPI; 2001-007495/01.
XX N-PSDB; AAC99752.
XX
XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases.
XX
XX
XX Claim 4; Page 296; 352pp; English.
XX
XX The present sequence is a polypeptide which is expressed in mammalian
CC skin cells. The polypeptide is useful for stimulating keratinocyte growth
CC and motility, inhibiting the growth of cancer cells, modulating
CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,
CC modulating skin inflammation, stimulating the growth of epithelial cells,
CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to
CC leukocytes, and treating inflammatory disease, cancer and neurological
CC diseases. The polynucleotide can be used as a marker, in the
CC identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns
XX
XX SQ Sequence 85 AA;

Query Match 1.4%; Score 17; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNARFGS 544
Db 28 NGTLKDSHSYQNARFGS 44

RESULT 45
AB872250
ID ABB72250 standard; protein; 85 AA.
XX
XX AC ABB72250;
XX
XX DT 04-APR-2002 (first entry)
XX
XX DE Rat protein isolated from skin cells SEQ ID NO: 390.
XX
XX DE Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
XX KW developmental defect; inflammatory disease; dermatological; vulnary;
XX KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
XX
XX OS Rattus sp.
XX
XX PN WO200190357-A1.
XX
XX PD 29-NOV-2001.
XX
XX PF 24-MAY-2001; 2001WO-NZ000099.
XX
XX PR 24-MAY-2000; 2000US-0206650P.
XX PR 25-JUL-2000; 2000US-0221232P.
XX
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
XX PI Kumble KD;
XX
XX DR WPI; 2002-122020/16.
XX
XX PT New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.
XX
XX PS Example 2; Page 245; 466pp; English.
XX

CC The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
XX polypeptide of the invention
XX SQ Sequence 85 AA;

Query Match 1.4%; Score 17; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNARFGS 544
Db 28 NGTLKDSHSYQNARFGS 44

RESULT 46
AAU19634
ID AAU19634 standard; protein; 148 AA.
XX
XX AC AAU19634;
XX
XX DT 04-DEC-2001 (first entry)
XX
XX DE Human novel extracellular matrix protein, Seq ID No 284.
XX
XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
XX KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
XX KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
XX KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
XX KW cancers; hyperproliferative disorder; breast neoplasia; melanoma;
XX KW Sezary syndrome; Gaucher's disease; neurological diseases;
XX KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
XX KW cardiac arrest; tachycardia; angina; infection; corneal infections;
XX KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
XX OS Homo sapiens.
XX
XX PN WO200153368-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001348.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184564P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216680P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218250P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.

PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226275P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228934P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0246471P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246529P.
PR	08-NOV-2000;	2000US-0246530P.
PR	08-NOV-2000;	2000US-0246531P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246533P.
PR	08-NOV-2000;	2000US-0246534P.
PR	08-NOV-2000;	2000US-0246535P.
PR	08-NOV-2000;	2000US-0246536P.
PR	08-NOV-2000;	2000US-0246537P.
PR	08-NOV-2000;	2000US-0246538P.
PR	08-NOV-2000;	2000US-0246539P.
PR	08-NOV-2000;	2000US-0246540P.
PR	08-NOV-2000;	2000US-0246541P.
PR	08-NOV-2000;	2000US-0246542P.
PR	08-NOV-2000;	2000US-0246543P.
PR	08-NOV-2000;	2000US-0246544P.
PR	08-NOV-2000;	2000US-0246545P.
PR	08-NOV-2000;	2000US-0246546P.
PR	08-NOV-2000;	2000US-0246547P.
PR	08-NOV-200	

CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC (e.g. corneal infections). Other uses include wound healing, maintenance
 CC of organs before transplantation, support of cell culture of primary

Query Match 1.38; Score 15; DB 4; Length 148;
 Best Local Similarity 100.0%; Pred. NO. 6.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYFW 178
 |||||
 Db 47 DIVVLGSGNSIYFW 61

RESULT 47

AAU19794
 ID AAU19794 standard; protein; 148 AA.

XX AC AAU19794;

XX DT 04-DEC-2001 (first entry)

XX DE Human novel extracellular matrix protein, Seq ID No 444.

XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
 KW antianemic; antirheumatic; antisclerotic; cardiant; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
 KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
 KW Sezary syndrome; Gaucher's disease; neurological diseases;
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
 KW Cardiac arrest; tachycardia; angina; infection; corneal infections;
 KW wound healing; immunogen; Gene therapy; antisense; food additive.

XX OS Homo sapiens.

XX PN WO200155368-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001348.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 30-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0215647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234233P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235844P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0248526P.
 PR 08-NOV-2000; 2000US-0248527P.
 PR 08-NOV-2000; 2000US-0248528P.
 PR 08-NOV-2000; 2000US-0248532P.
 PR 08-NOV-2000; 2000US-0248609P.
 PR 08-NOV-2000; 2000US-0248610P.
 PR 08-NOV-2000; 2000US-0248611P.

PR	08-NOV-2000;	2000US-0246613P.	Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
PR	17-NOV-2000;	2000US-0249207P.	Qy	164	DIVIVLDGNSIYPW	178						
PR	17-NOV-2000;	2000US-0249208P.										
PR	17-NOV-2000;	2000US-0249208P.										
PR	17-NOV-2000;	2000US-0249210P.	Db	47	DIVIVLDGNSIYPW	61						
PR	17-NOV-2000;	2000US-0249211P.										
PR	17-NOV-2000;	2000US-0249212P.										
PR	17-NOV-2000;	2000US-0249213P.										
PR	17-NOV-2000;	2000US-0249214P.										
PR	17-NOV-2000;	2000US-0249215P.										
PR	17-NOV-2000;	2000US-0249216P.										
PR	17-NOV-2000;	2000US-0249217P.										
PR	17-NOV-2000;	2000US-0249218P.										
PR	17-NOV-2000;	2000US-0249244P.										
PR	17-NOV-2000;	2000US-0249245P.										
PR	17-NOV-2000;	2000US-0249264P.										
PR	17-NOV-2000;	2000US-0249265P.										
PR	17-NOV-2000;	2000US-0249297P.										
PR	17-NOV-2000;	2000US-0249299P.										
PR	17-NOV-2000;	2000US-0249300P.										
PR	01-DEC-2000;	2000US-0250160P.										
PR	01-DEC-2000;	2000US-0250391P.										
PR	05-DEC-2000;	2000US-0251030P.										
PR	05-DEC-2000;	2000US-0251988P.										
PR	05-DEC-2000;	2000US-0256719P.										
PR	06-DEC-2000;	2000US-0251479P.										
PR	08-DEC-2000;	2000US-0251856P.										
PR	08-DEC-2000;	2000US-0251868P.										
PR	08-DEC-2000;	2000US-0251869P.										
PR	08-DEC-2000;	2000US-0251989P.										
PR	08-DEC-2000;	2000US-0251990P.										
PR	11-DEC-2000;	2000US-0254097P.										
PR	05-JAN-2001;	2001US-0259678P.										
XX												
PA	(HUMA-) HUMAN GENOME SCI INC.											
XX												
PI	Rosen CA, Barash SC, Ruben SM;											
XX												
XX												
DR	WTI; 2001-455572/50.											
DR	N-PSDB; AAS31365.											
XX												
PT	Nucleic acid molecules encoding human secreted extracellular matrix											
PT	proteins, used in preventing, treating or ameliorating a disorder, e.g.											
PT	Alzheimer's and Parkinson's diseases and cancers.											
XX												
PS	Claim 11; SEQ ID NO 444; 577pp; English.											
XX												
CC	The invention relates to isolated nucleic acid molecules encoding novel											
CC	human secreted extracellular matrix proteins (SPs). The polynucleotides											
CC	and proteins are used to prevent, treat a medical condition in e.g.											
CC	humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For											
CC	example, disorders associated with decreased expression of SPs. The SP											
CC	polynucleotide or a vector expressing them may be administered to treat											
CC	diseases by gene therapy. Antisense molecules may be administered to down											
CC	regulate expression of SPs by binding with the cells own genes and											
CC	preventing their expression. The polynucleotides may also be used as DNA											
CC	probes in diagnostic assays. The SPs may also be used as antigens to											
CC	produce antibodies and to identify modulators (agonists and antagonists)											
CC	of the SPs. The anti-(SP) antibodies and antagonists may also be used to											
CC	down regulate expression and activity of SP and as diagnostic agents for											
CC	detecting the presence of SPs in samples. The disorders include for											
CC	example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency											
CC	virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),											
CC	cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of											
CC	the breast or liver, Sezary syndrome and Gaucher's disease), neurological											
CC	diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-											
CC	/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),											
CC	infections caused by bacteria, viruses and fungi and ocular disorders											
CC	(e.g. corneal infections). Other uses include wound healing, maintenance											
CC	of organs before transplantation, support of cell culture of primary											

Query Match Similarity 1.3%; Score 15; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;

```

PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-470713/50.
XX N-PSDB; ABQ66529.
XX
XX New nucleic acid encoding human proteins, useful for diagnosis, treatment
XX and prevention of e.g. osteoporosis, also related polypeptides and
XX antibodies.
XX
XX Claim 11; SEQ ID NO 284; 235pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
XX (ABP47846-ABP4810) useful for preventing, treating or ameliorating
XX medical conditions e.g. By protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html?DocID=99909764870
XX
XX Sequence 148 AA;
XX
XX Query Match 1.3%; Score 15; DB 5; Length 148;
XX Best Local Similarity 100.0%; Pred. No. 6.9e-06;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 164 DIVIVLDGNSIYPW 178
XX Db 47 DIVIVLDGNSIYPW 61
XX
XX RESULT 49
XX ID ABP48014
XX AB 164 DIVIVLDGNSIYPW 178
XX AC ABP48014;
XX XX
XX XX 23-AUG-2002 (first entry)
XX DE Human polypeptide SEQ ID NO 444.
XX
XX Human; neutrotropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX

```

```

OS Homo sapiens.
XX
XX US2002042386-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-00764870.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 28-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214896P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220984P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 22-AUG-2000; 2000US-0226868P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 08-SEP-2000; 2000US-0231413P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0235834P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 29-SEP-2000; 2000US-0236370P.
XX 02-OCT-2000; 2000US-0236802P.
XX 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
XX 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
XX 13-OCT-2000; 2000US-0239935P.
XX 20-OCT-2000; 2000US-0240960P.
XX 20-OCT-2000; 2000US-0241785P.
XX 20-OCT-2000; 2000US-0241809P.
XX 01-NOV-2000; 2000US-0244617P.
XX 17-NOV-2000; 2000US-0249299P.
XX 08-DEC-2000; 2000US-0251856P.
XX 08-DEC-2000; 2000US-0251868P.
XX 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-470713/50.
XX N-PSDB; ABQ66689.
XX
XX New nucleic acid encoding human proteins, useful for diagnosis, treatment
XX and prevention of e.g. osteoporosis, also related polypeptides and
XX antibodies.
XX
XX Claim 11; SEQ ID NO 444; 235pp + Sequence Listing; English.
XX

```

XX CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins (ABP47846-ABP48110) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease, multiple sclerosis, thrombotic thrombocytopenic purpura; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=99903764870

XX SQ Sequence 148 AA;

Query Match 1.3%; Score 15; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPM 178
DB 47 DIVIVLDGNSIYPM 61
|||||

RESULT 50
ADCI0816
ID ADC10816 standard; protein; 148 AA.
AC ADC10816;
XX
XX
DT 18-DEC-2003 (first entry)
XX
DE Human extracellular matrix protein from gene 9.
XX
XX Extracellular matrix protein; cytosolic; antibacterial; virucide;
XX neuroprotective; gynaecological; gastrointestinal; cardiac;
XX cardiovascular; nephrotropic; antiinflammatory; muscular; Gen;
XX respiratory; Gen; immunosuppressive; cerebroprotective; vasotropic;
XX neotropic; anti-allergic; cancer; bacterial infection; viral infection;
XX neural disorder; immune system disorder; blood disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX inflammatory disorder; proliferative disorder; Human.
XX
XX Homo sapiens.
XX
XX US2003059875-A1.
XX
XX 27-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-01981123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214896P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244825P.
PR 01-NOV-2000; 2000US-0244826P.

PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246529P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251031P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256179P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254037P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764870.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PA
 XX Rosen CA, Ruben SM, Barash SC;
 XX PI
 XX WPI; 2003-743755/70.
 XX DR N-PSDB; ADCI0531.
 XX
 PT New isolated nucleic acids and polypeptides, useful for diagnosing,
 PT treating, and/or preventing disorders, such as cancer, infections,
 PT cardiovascular and inflammatory diseases.
 XX
 PS Claim 11; SEQ ID NO 284; 235pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (cDNA)
 CC encoding a human extracellular matrix protein, representing one of 161
 CC novel genes. Also included are recombinant vectors, host cells
 CC (expressing the protein), the extracellular matrix proteins (including
 CC their fragments, epitopes and homologues), an isolated antibody that
 CC binds specifically to the protein, diagnosing a pathological condition or
 CC susceptibility to a pathological condition (comprising determining the
 CC presence or absence of a mutation in the nucleic acid and diagnosing a
 CC condition based on the presence or absence of the mutation), diagnosing a
 CC pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the
 CC protein in a biological sample and diagnosing a condition based on the
 CC presence or amount of expression of the protein), preventing, treating or
 CC ameliorating a medical condition by administering the nucleic acid or
 CC protein to a mammalian subject, identifying a binding partner to the
 CC protein, the gene corresponding to the cDNA sequence, and identifying an
 CC activity in a biological assay (comprising expressing the nucleic acid in
 CC a cell, isolating the supernatant, detecting an activity in a biological
 CC assay and identifying the protein in the supernatant having the
 CC activity). The nucleic acids and proteins display the following
 CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,
 CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,
 CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,
 CC

Query Match 1.3%; Score 15; DB 7; Length 148;
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVDGSNSIYPW 178
 |||||
 DB 47 DIVIVDGSNSIYPW 61

RESULT 51

ADC10976
 ID ADC10976 standard; protein; 148 AA.

AC ADC10976;

DT 18-DEC-2003 (first entry)

DE Human protein from extracellular matrix gene 9 #2.

XX Extracellular matrix protein; cytostatic; antibacterial; virucide;
 KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
 KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
 KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
 KW notropic; antiallergic; cancer; bacterial infection; viral infection;
 KW neural disorder; immune system disorder; blood disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW inflammatory disorder; proliferative disorder; Human.

XX Homo sapiens.

XX US2003059875-A1.

XX 27-MAR-2003.

XX 19-APR-2002; 2002US-00125540.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216547P.

XX 07-JUL-2000; 2000US-0216800P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0228868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 13-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235936P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764870.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-743765/70.
XX N-PSDB; ADC10711.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating, and/or preventing disorders, such as cancer, infections,
XX cardiovascular and inflammatory diseases.
XX
XX Claim 11; SEQ ID NO 444; 235pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA)
XX encoding a human extracellular matrix protein, representing one of 161
XX novel genes. Also included are recombinant vectors, host cells
XX (expressing the protein), the extracellular matrix proteins (including
XX their fragments, epitopes and homologues), an isolated antibody that
XX binds specifically to the protein, diagnosing a pathological condition or
XX susceptibility to a pathological condition (comprising determining the
XX presence or absence of a mutation in the nucleic acid and diagnosing a
XX condition based on the presence or absence of the mutation), diagnosing a
XX pathological condition or susceptibility to a pathological condition
XX (comprising determining the presence or amount of expression of the
XX protein in a biological sample and diagnosing a condition based on the
XX presence or amount of expression of the protein), preventing, treating or
XX ameliorating a medical condition by administering the nucleic acid or
XX protein to a mammalian subject, identifying a binding partner to the
XX protein, the gene corresponding to the cDNA sequence, and identifying an
XX activity in a biological assay (comprising expressing the nucleic acid in

CC a cell, isolating the supernatant, detecting an activity in a biological
 CC assay and identifying the protein in the supernatant having the
 CC activity). The nucleic acids and proteins display the following
 CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,
 CC Gynaecological, Gastrointestinal-Gen, Cardiac, Cardiovascular-Gen,
 CC Nephrologic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 1.3%; Score 15; DB 7; Length 148;
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178
 Db 47 DIVVLGSGNSIYPW 51

RESULT 52
 AAE33535
 ID AAE33535 standard; protein; 192 AA.

XX AC AAE33535;

XX DT 23-OCT-2003 (revised)

XX DT 02-APR-2003 (first entry)

XX DE Rat-human alpha 1-I domain fusion protein, RdeltaH.

XX KW Human; very late activation antigen; VLA-1; beta1 containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia; fusion protein; rat.

XX OS Homo sapiens.

XX OS Rattus sp.

XX OS Chimeric.

XX PN WO200283854-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011521.

XX PR 13-APR-2001; 2001US-0283794P.

XX PR 06-JUL-2001; 2001US-0303689P.

XX PA (BIOJ) BIOGEN INC.

XX PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX DR WPI; 2003-093009/08.

XX PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.

XX PS Example 24; Page 93; 248pp; English.

XX CC The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal

CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
 CC periarteritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC rat-human alpha 1-I domain fusion protein, RdeltaH. This sequence is used
 CC in the exemplification of the invention. (Updated on 23-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;

Best Local Similarity 100.0%; Pred. No. 8.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178

Db 4 DIVVLGSGNSIYPW 18

RESULT 53

AAE33536

ID AAE33536 standard; protein; 192 AA.

XX AC AAE33536;

XX DT 02-APR-2003 (first entry)

XX DT Rat alpha 1-I domain protein #1.

XX KW Rat; very late activation antigen; VLA-1; beta1 containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia, alpha 1-I protein.

XX OS Rattus sp.

XX PN WO200283854-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011521.

XX PR 13-APR-2001; 2001US-0283794P.

XX PR 06-JUL-2001; 2001US-0303689P.

XX PA (BIOJ) BIOGEN INC.

XX PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX DR WPI; 2003-093009/08.

XX PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.

PT fibrosis.

XX PS Example 24; Page 93-94; 248pp; English.

XX

CC The present invention relates to novel antibodies that specifically bind to very late activation (VLA-1; beta1 containing integrins) antigens and methods of using these antibodies to treat immunological disorders. The anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders such as skin related conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal proliferation of hair follicle cells), fibrosis (e.g. kidney or lung fibrosis), allergic rhinitis, respiratory distress syndrome, asthma, bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease, gastritis, irritable bowel syndrome, colitis and colorectal cancer), vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia, periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis, autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid arthritis, systemic lupus erythematosus and multiple sclerosis), renal failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis, polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or immediate hypersensitivity), graft and transplant rejections, graft versus host disease, conjunctivitis, swelling occurring after injury, myocardial ischaemia or endotoxin shock syndrome. The present sequence is rat 1-I domain protein. This sequence is used in the exemplification of the invention

XX SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
Db 4 DIVIVLDGNSIYPW 18

RESULT 54
AAE33537
ID AAE33537 standard; protein; 192 AA.

AC AAE33537;

XX 02-APR-2003 (first entry)

DT Human alpha 1-I domain protein #1.

DE

XX Human; very late activation antigen; VLA-1; beta1 containing integrin; immunological disorder; inflammatory disorder; skin related condition; psoriasis; eczema; burn; dermatitis; respiratory distress syndrome; fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis; fever; migraine headache; inflammatory bowel disease; Crohn's disease; irritable bowel syndrome; colitis; colorectal cancer; vascular disease; atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa; gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease; osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome; renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis; hypersensitivity; graft rejection; transplant rejection; conjunctivitis; graft versus host disease; myocardial ischaemia; alpha 1-I protein.

XX OS Homo sapiens.

XX WO200283854-A2.

PN 24-OCT-2002.

XX

XX 12-APR-2002; 2002WO-US011521.

PF

XX 13-APR-2001; 2001US-0283794P.

PR

XX 06-JUL-2001; 2001US-0303689P.

PR

XX

PA (BIOJ) BIOGEN INC.

XX LYNE PD, Garber EA, Saldanha JW, Karpusas M;

XX WPI; 2003-093009/08.

DR

XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders, e.g. psoriasis, eczema, burns, dermatitis, and abnormal proliferation of hair follicle cells or fibrosis.

XX Example 24; Page 94; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind to very late activation (VLA-1; beta1 containing integrins) antigens and methods of using these antibodies to treat immunological disorders. The anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders such as skin related conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal proliferation of hair follicle cells), fibrosis (e.g. kidney or lung fibrosis), allergic rhinitis, respiratory distress syndrome, asthma, bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease, gastritis, irritable bowel syndrome, colitis and colorectal cancer), vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia, periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis, autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid arthritis, systemic lupus erythematosus and multiple sclerosis), renal failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis, polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or immediate hypersensitivity), graft and transplant rejections, graft versus host disease, conjunctivitis, swelling occurring after injury, myocardial ischaemia or endotoxin shock syndrome. The present sequence is human 1-I domain protein. This sequence is used in the exemplification of the invention

XX SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
Db 4 DIVIVLDGNSIYPW 18

RESULT 55
AAU76851
ID AAU76851 standard; protein; 195 AA.

XX AAU76851;

XX 21-MAY-2002 (first entry)

DT Human integrin alpha subunit Alpha 1 (CD49a) A domain.

DE

XX Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit; A-like domain; inflammatory disorder; skeletal muscle injury; restenosis; ischaemia-reperfusion injury; immune complex; parasitic disease; Alpha 1; antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.

XX OS Homo sapiens.

XX WO200209737-A1.

PN 07-FEB-2002.

XX

XX 31-JUL-2001; 2001WO-US023957.

PF

XX 31-JUL-2000; 2000US-0221950P.

PR

XX 11-JAN-2001; 2001US-00758493.

PR

XX 13-MAR-2001; 2001US-00805354.

```

XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI
XX PI Arnaut AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Example 2; Fig 5; 55pp; English.
XX CC
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 1 (CD49a) A domain
XX SQ Sequence 195 AA;

Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DQ 6 DIVIVLDGNSIYPW 20
|||||
|||||

RESULT 56
AAU76860
ID AAU76860 standard; protein; 195 AA.
XX AC AAU76860;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 1 (CD49a) variant A domain.
XX KW Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
XX KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy;
XX KW mutin; Alpha 1.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 193
XX FT /note= "Wild-type Ile substituted by any other amino
XX FT acid"
XX PN WC200209737-A1.
XX XX
XX PD 07-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US023957.
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI
XX PI Arnaut AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Claim 45; Page; 55pp; English.
XX CC
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 1 (CD49a) variant A domain. Note: This
XX CC variant sequence is not featured in the specification but has been
XX CC derived from the wild-type protein shown in AAU76851
XX SQ Sequence 195 AA;

Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DQ 6 DIVIVLDGNSIYPW 20
|||||
|||||

RESULT 57
AAB50041
ID AAB50041 standard; protein; 214 AA.
XX AC AAB50041;
XX DT 14-MAR-2001 (first entry)
XX DE Rat alaphal integrin domain.
XX KW Rat; alaphal integrin domain; alaphal beta function blocking antibody;
XX KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
XX KW bronchitis; headache; antipyretic; fever; gastrointestinal;
XX KW vascular disease; autoimmune; respiratory distress syndrome;
XX KW endotoxin shock syndrome; atherosclerosis.
XX OS Rattus sp.
XX PN WC2000072881-A1.
XX XX
XX PD 07-DEC-2000.
XX PF 01-JUN-2000; 2000WO-US015004.
XX PR 01-JUN-1999; 99US-0137038P.
XX PR 29-FEB-2000; 2000US-0185336P.
XX PA (BIOJ ) BIOGEN INC.
XX PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;
XX XX

```



```

DR WPI; 2001-061448/07.
XX
PT Use of blocking monoclonal antibody capable of binding to an epitope of
PT VLA-1 for treating inflammatory disorders, in particular arthritis.
XX
PS Example 14; Fig 14; 60pp; English.
XX
CC The present invention relates to the use of an alphabeta1 function
CC blocking antibody capable of binding an epitope of human alpha1 integrin
CC domain (VLA-1; see AAB50043). The antibody of the present invention is
CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
CC treatment of pain and headaches, or as an antipyretic, for the treatment
CC of fever, gastrointestinal conditions such as inflammatory bowel
CC diseases, Crohn's disease, gastritis and vascular diseases, migraine
CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's
CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple
CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,
CC allergic rhinitis, respiratory distress syndrome, endotoxin shock
CC syndrome and atherosclerosis. The present sequence is the rat alpha1
CC integrin domain
XX
SQ Sequence 214 AA;
    Query Match 1.3%; Score 15; DB 4; Length 214;
    Best Local Similarity 100.0%; Pred. No. 9.5e-06;
    Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 22 DIVIVLDGNSIYPW 36

RESULT 58
AAB50042
ID AAB50042 standard; protein; 214 AA.
AC AAB50042;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human alpha1 integrin domain.
XX
KW Human; alpha1 integrin domain; alphabeta1 function blocking antibody;
KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
KW bronchitis; headache; antipyretic; fever; gastrointestinal;
KW vascular disease; autoimmune; respiratory distress syndrome;
KW endotoxin shock syndrome; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200072881-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US015004.
XX
PR 01-JUN-1999; 99US-0137038P.
PR 29-FEB-2000; 2000US-0185336P.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;
XX
XX WPI; 2001-061448/07.
XX
PT Use of blocking monoclonal antibody capable of binding to an epitope of
PT VLA-1 for treating inflammatory disorders, in particular arthritis.
XX
XX Example 14; Fig 15; 60pp; English.
XX
CC The present invention relates to the use of an alphabeta1 function

```

```

CC blocking antibody capable of binding an epitope of human alpha1 integrin
CC domain (VLA-1; see AAB50043). The antibody of the present invention is
CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
CC treatment of pain and headaches, or as an antipyretic, for the treatment
CC of fever, gastrointestinal conditions such as inflammatory bowel
CC diseases, Crohn's disease, gastritis and vascular diseases, migraine
CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's
CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple
CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,
CC allergic rhinitis, respiratory distress syndrome, endotoxin shock
CC syndrome and atherosclerosis. The present sequence is the human alpha1
CC integrin domain
XX
SQ Sequence 214 AA;
    Query Match 1.3%; Score 15; DB 4; Length 214;
    Best Local Similarity 100.0%; Pred. No. 9.5e-06;
    Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 22 DIVIVLDGNSIYPW 36

RESULT 59
AAB33539
ID AAB33539 standard; protein; 214 AA.
XX
AC AAB33539;
XX
DT 02-APR-2003 (first entry)
XX
DE Rat alpha 1-I domain protein #2.
XX
KW Rat; very late activation antigen; VLA-1; beta1 containing integrin;
KW immunological disorder; inflammatory disorder; skin related condition;
KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
KW fibrosis; allergic rhinitis; asthma; bronchitis; tendinitis; bursitis;
KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
KW atherosclerosis; thyroiditis; rheumatic fever; autoimmune disease;
KW gastritis; Hodgkin's disease; aplastic anaemia; periarthritis nodosa;
KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
KW Graft versus host disease; myocardial ischaemia; alpha 1-I protein.
XX
OS Rattus sp.
XX
PN WO200283854-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US011521.
XX
PR 13-APR-2001; 2001US-0283794P.
PR 06-JUL-2001; 2001US-0303689P.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;
XX
XX WPI; 2003-093009/08.
XX
PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
PT fibrosis.
XX
XX Example 18; Fig 11A; 248pp; English.

```

XX The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC periarthritis nodosa, Hodgkin's disease, rheumatoid arthritis, aplastic
 CC autoimmune diseases (e.g. atherosclerosis), thyroiditis, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC rat 1-I domain protein. This sequence is used in the exemplification of
 CC the invention

XX Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYFW 178
 |||||
 Db 22 DIVVLGSGNSIYFW 36

RESULT 60
 AAE33540
 ID AAE33540 standard; protein; 214 AA.

XX AAE33540;
 DT 02-APR-2003 (first entry)
 DE Human alpha 1-I domain protein #2.

XX Human; very late activation antigen; VLA-1; beta1 containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.

XX Homo sapiens.
 OS WO200283854-A2.
 XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011521.
 XX 13-APR-2001; 2001US-0283794P.
 XX 06-JUL-2001; 2001US-0303689P.
 XX (BIO) BIOGEN INC.
 XX Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX WPI; 2003-093009/08.
 XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.
 XX Example 19; Fig 12; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
 CC periarthritis nodosa, Hodgkin's disease, rheumatoid arthritis, myasthenia gravis, rheumatoid
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC human 1-I domain protein. This sequence is used in the exemplification of
 CC the invention

XX Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYFW 178
 |||||
 Db 22 DIVVLGSGNSIYFW 36

RESULT 61
 ADEB6652
 ID ADEB6652 standard; protein; 1151 AA.

XX ADEB6652;
 DT 29-JAN-2004 (first entry)
 DE Human integrin alpha 1 subunit.

XX human; secreted protein; cancer; liver disorder; hepatitis;
 KW neural disorder; Alzheimer's disease; integrin alpha 1 subunit.
 XX Homo sapiens.
 OS US2003129685-A1.
 XX 10-JUL-2003.
 XX 18-APR-2001; 2001US-00836353.
 XX 28-OCT-1998; 98US-0105971P.
 XX 27-OCT-1999; 99WO-US025031.
 XX 19-APR-2000; 2000US-0198407P.
 XX (NIJ) NI J.
 XX (YOUNG) YOUNG P E.
 XX (KENDRICK) KENDRICK J J.
 XX (OLSEN) OLSEN H S.

PA (MOORE/) MOORE P A.
PA (WEIY/) WEI Y.
PA (GREENE/) GREENE J M.
PA (RUBEN/) RUBEN S M.
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
XX Ruben SM;
XX WPI; 2004-0203335/02.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX PT
XX PS Disclosure; SEQ ID NO 103; 380pp; English.
XX CC The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of an integrin alpha 1 subunit.
XX CC
XX SQ Sequence 1151 AA;
Query Match 1.3%; Score 15; DB 8; Length 1151;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVLDGNSIYPM 178
DB 144 DIVIVLDGNSIYPM 158
RESULT 62
ABB90759
ID ABB90759 standard; protein; 1179 AA.
XX AC ABB90759;
XX DT 30-MAY-2002 (first entry)
XX DE Human Tumour Endothelial Marker polypeptide SEQ ID NO 250.
XX KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
KW normal endothelial marker; pan-endothelial marker; immunostimulant;
KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW psoriasis.
XX OS Homo sapiens.
XX PN WO200210217-A2.
XX PD 07-FEB-2002.
XX PF 01-AUG-2001; 2001WO-US024031.
XX PR 02-AUG-2000; 2000US-0222599P.
XX PR 11-AUG-2000; 2000US-0224360P.
XX PR 11-APR-2001; 2001US-0282850P.
XX FA (UKJO) UNIV JOHNS HOPKINS.
XX ST Croix B, Kinzler KW, Vogelstein B;
XX WPI; 2002-291856/33.
XX DR N-PSDB; ABL92113.
XX PT An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth.

XX Claim 35; Page 242-245; 331pp; English.
XX The invention relates to an isolated molecule comprising an antibody
XX variable region which specifically binds to an extracellular domain of a
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects
CC bearing a vascularised tumour, polycystic kidney disease, diabetic
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)
CC are disclosed, as are marker oligonucleotide sequences: tumour
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
CC (PEM) ABL91903-ABL91995
XX SQ Sequence 1179 AA;
Query Match 1.3%; Score 15; DB 5; Length 1179;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVLDGNSIYPM 178
DB 172 DIVIVLDGNSIYPM 186
RESULT 63
ABP64915
ID ABP64915 standard; protein; 1179 AA.
XX AC ABP64915;
XX DT 25-FEB-2003 (first entry)
XX DE Human protein SEQ ID 575.
XX KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective.
XX OS Homo sapiens.
XX PN WO200259260-A2.
XX PD 01-AUG-2002.
XX PF 16-NOV-2001; 2001WO-US042950..
XX PR 17-NOV-2000; 2000US-00714936.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
XX Ren F, Xue RJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-590824/63.
XX DR N-PSDB; ABQ99501.
XX PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX PS Claim 20; SEQ ID NO 575; 394pp; English.
XX CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in

CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotential or pluripotential state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 5; Length 1179;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 172 DIVVLGSGNSIYPW 186

RESULT 64
 ABUS4466
 ID ABUS4466 standard; protein; 1179 AA.

XX AC ABUS4466;

XX DT 12-MAR-2003 (first entry)

XX DE Human tumour endothelial marker TEM 30.

XX KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KW Tumour endothelial marker; normal endothelial marker; PEM;
 KW Pan-endothelial marker; polycystic kidney disease; psoriasis;
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KW neovascularization; immune response; cytostatic; antidiabetic;
 KW ophthalmological; anti-rheumatic; anti-arthritis; antipsoriatic.

XX OS Homo sapiens.

XX PN WO200283874-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US008253.

XX PR 11-APR-2001; 2001US-0282850P.

XX PR 06-FEB-2002; 2002US-0354262P.

XX PA (UYJO) UNIV JOHNS HOPKINS.

XX PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX DR WPI; 2003-093016/08.

XX DR N-PSDB; ABX72038.

XX PT New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,

PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.

XX PS Disclosure; Page 267-269; 374pp; English.

XX CC The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for
 CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a human TEM or NEM protein of the invention
 XX
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 6; Length 1179;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 172 DIVVLGSGNSIYPW 186

RESULT 65
 ABB90788
 ID ABB90788 standard; protein; 1180 AA.

XX AC ABB90788;

XX DT 30-MAY-2002 (first entry)

XX DE Rat Tumour Endothelial Marker polypeptide SEQ ID NO 307.

XX KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis.

XX OS Rattus sp.

XX PN WO200210217-A2.

XX PD 07-FEB-2002.

XX PF 01-AUG-2001; 2001WO-US024031.

XX PR 02-AUG-2000; 2000US-0222599P.

XX PR 11-AUG-2000; 2000US-0224360P.

XX PR 11-APR-2001; 2001US-0282850P.

XX PA (UYJO) UNIV JOHNS HOPKINS.

XX PI St Croix B, Kinzler KW, Vogelstein B;

XX DR WPI; 2002-291856/33.

XX DR N-PSDB; ABL92141.

XX PT An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth.

XX PS Disclosure; Page 314-318; 331pp; English.

XX CC The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic
 CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
 CC genes and the encoded proteins (ABL92075-ABU92141 and ABB90721-ABB90789)
 CC are disclosed, as are marker oligonucleotide sequences: tumour
 CC endothelial markers (TEM) ABL91996-ABU92041 and ABL92143-ABU92191; normal
 CC endothelial markers (NEM) ABL92042-ABU92074; and pan-endothelial markers
 CC (PEM) ABL91903-ABL91995
 XX
 SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 5; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVVLGSGNSIYPW 178
 Db 172 DIVVLGSGNSIYPW 186
 |||||

RESULT 66
 ABU54495
 ID ABU54495 standard; protein; 1180 AA.
 XX
 AC ABU54495;

XX
 DT 12-MAR-2003 (first entry)

XX Mouse tumour endothelial marker mTEM 30.

XX Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KW Tumour endothelial marker; normal endothelial marker; PEM;
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KW neoangiogenesis; immune response; cytostatic; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX Mus sp.

XX WO200283874-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US008253.

XX 11-APR-2001; 2001US-0282850P.

XX 06-FEB-2002; 2002US-0354262P.

XX (CVJO) UNIV JOHNS HOPKINS.

XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2003-093016/08.

XX N-PSDB; ABX72066.

XX New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.

XX Disclosure; Page 356-358; 374pp; English.

XX The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM) normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for

CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a mouse TEM protein

SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 6; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVVLGSGNSIYPW 178
 Db 172 DIVVLGSGNSIYPW 186
 |||||

RESULT 67
 ADE63568
 ID ADE63568 standard; protein; 1180 AA.

XX
 AC ADE63568;

XX 29-JAN-2004 (first entry)

XX Rat Protein P18614, SEQ ID NO 9512.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P18614.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 7; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNSIYPW 178
 |||||
 Db 172 DIVVLGDSNSIYPW 186

RESULT 68
 AAB30928
 ID AAB30928 standard; peptide; 15 AA.

XX AC AAB30928;

XX DT 02-APR-2001 (first entry)

XX DE Antigenic peptide derived from a human alpha11 integrin chain.

XX KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
 XX KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
 XX KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
 XX KW osteoporosis; cartilage damage; bone damage; cartilage.

XX OS Homo sapiens.

XX PN WO200075187-A1.

XX PD 14-DEC-2000.

XX PF 31-MAY-2000; 2000WO-SF001135.

XX PR 03-JUN-1999; 99SF-00002056.

XX PA (ACTI-) ACTIVE BIOTECH AB.

XX PI Gullberg D;

XX DR WPI; 2001-071061/08.

XX PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit
 PT alpha 11 in association with subunit beta, useful for treating muscle
 PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX PS Disclosure; Page 17; 79pp; English.

XX CC The present sequence is derived from the cytoplasmic domain of the human
 CC integrin subunit, designated alpha11. The alpha11 polynucleotide and
 CC polypeptide are useful as markers of cell target molecules, such as
 CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
 CC derived cells or stem cells. They are also used for determining the
 CC differential-stage of cells during differentiation, development in
 CC pathological conditions, in tissue regeneration, in transplantation or in
 CC therapeutic and physiological repair of tissues. The pathological
 CC conditions involving subunit alpha11 are selected from damage of cells,
 CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
 CC osteoarthritis and osteoporosis, damage of cartilage and bone, and
 CC cartilage and bone diseases. The polypeptide is useful for detecting the
 CC formation of cartilage during embryonic development, for detecting
 CC physiological therapeutic repair of cartilage and muscle, for selection
 CC and analysis, or for sorting, isolating or purification of chondrocytes
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
 CC during transplantation of cartilage or chondrocytes during

CC transplantation of cartilage or chondrocytes, respectively, or of muscle
 CC or muscle cells during transplantation of muscle or muscle cells,
 CC respectively, and for studies of differentiation or chondrocytes or
 CC muscle cells

XX SQ Sequence 15 AA;

Query Match 1.2%; Score 14; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.1e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1175 RREPLDPTPKVLE 1188
 |||||

Db 2 RREPLDPTPKVLE 15

RESULT 69
 AAY07728
 ID AAY07728 standard; protein; 1183 AA.

XX AC AAY07728;

XX DT 01-JUL-1999 (first entry)

XX DE Armenian hamster alpha-1 integrin subunit protein.

XX KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;
 XX KW integrin cell surface receptor; capillary; blood vessel; hamster;
 XX KW alpha-1 subunit; alpha-2 subunit.

XX OS Cricetulus migratorius.

XX PN WO9916465-A1.

XX PD 08-APR-1999.

XX PF 30-SEP-1997; 97WO-US017485.

XX PR 30-SEP-1997; 97WO-US017485.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (SENG/) SENG D R.

XX PA (DETM/) DETMAR M.

XX PA (CLAF/) CLAFFEY K P.

XX PI Senger DR, Detmar M, Claffey KP;

XX DR WPI; 1999-254930/21.

XX PT Inhibition of tumor angiogenesis through interaction of vascular
 PT endothelial growth factor and integrin cell surface receptors.

XX PS Disclosure; Fig 2A-C; 64pp; English.

XX CC This invention describes a novel method for the inhibition of tumour
 CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and
 CC integrin cell surface receptors expressed in vasculature of living
 CC subjects. The method inhibits new capillary and new blood vessel
 CC formation both within a tumour mass itself as well as in the immediately
 CC adjacent blood vasculature surrounding the perimeter of the tumour mass.
 CC interaction and dependence upon VEGF to induce specific integrin
 CC heterodimers in tumour angiogenesis provides a novel method for
 CC inhibiting tumour angiogenesis, and unlike prior art relies on the
 CC specific inter-relationship of VEGF and integrins, rather than
 CC concentrating solely on one specific class of protein

XX SQ Sequence 1183 AA;

Query Match 1.2%; Score 14; DB 2; Length 1183;
 Best Local Similarity 100.0%; Pred. No. 0.00042;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNSIYP 177

Db 174 DIVVLDGNSIYPW 187
|||||
RESULT 70
AAU76862
ID AAU76862 standard; protein; 195 AA.
XX AC AAU76862;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 10 variant A domain.
XX KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
XX KW antiinflammatory; vasotropic; antiparasitic; vulnerary; gene therapy;
XX KW mutin.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT FT Misc-difference 193
FT FT /note= "Wild-type Ile substituted by any other amino
FT FT acid"
XX PN WO200209737-A1.
XX XX
XX PD 07-FEB-2002.
XX XX
XX PF 31-JUL-2001; 2001WO-US023957.
XX XX
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.
XX XX
XX PA (GEO) GEN HOSPITAL CORP.
XX PI Arnaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Claim 53; Page; 55pp; English.
XX XX
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 10 variant A domain. Note: This variant
XX CC sequence is not featured in the specification but has been derived from
XX CC the wild-type protein shown in AAU76853
XX SQ Sequence 195 AA;
Query Match 1.1%; Score 13; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178
|||||
Db 8 VIVLDGNSIYPW 20
RESULT 71
AAU76853
ID AAU76853 standard; protein; 195 AA.
XX AC AAU76853;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 10 A domain.
XX KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease;
XX KW antiinflammatory; vasotropic; antiparasitic; vulnerary; gene therapy.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT FT Misc-difference 193
FT FT /note= "Wild-type Ile substituted by any other amino
FT FT acid"
XX PN WO200209737-A1.
XX XX
XX PD 07-FEB-2002.
XX XX
XX PF 31-JUL-2001; 2001WO-US023957.
XX XX
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.
XX XX
XX PA (GEO) GEN HOSPITAL CORP.
XX PI Arnaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Example 2; Fig 5; 55pp; English.
XX XX
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 10 A domain
XX SQ Sequence 195 AA;
Query Match 1.1%; Score 13; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AAY32243
 ID AAY32243 standard; protein; 1132 AA.
 AC AAY32243;
 XX 15-FEB-2000 (first entry)
 DT DT
 DE Human integrin subunit alpha-10 splice variant.
 DE
 XX Integrin alpha-10; ISa10; human; trauma; rheumatoid arthritis;
 XX osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;
 XX therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker;
 XX splice variant.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PH Peptide 1..22
 FT /note= "signal peptide"
 FT Protein 23..1132
 FT /note= "mature protein"
 XX
 XX WO9951639-A1.
 XX
 XX 14-OCT-1999.
 PD
 XX 31-MAR-1999; 99WO-SF000544.
 XX
 XX 02-APR-1998; 98SE-00001164.
 PR
 XX 28-JAN-1999; 99SE-00000319.
 XX
 XX (ACTI-) ACTIVE BIOTECH AB.
 PA
 XX Lundgren-Akerlund E;
 PI
 XX WPI; 2000-052639/04.
 DR
 XX N-PSDB; AAZ34720.
 DR
 XX New isolated integrin subunit alpha-10, used as a marker or target
 PT molecule for cells during development, regeneration and pathological
 PT conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or
 PT inflammation.
 XX
 XX Claim 1; Page 43-48; 90pp; English.
 PS
 XX This sequence represents a splice variant of novel human chondrocyte
 CC integrin subunit alpha-10 (ISa10). It is identical to ISa10 (see
 CC AAY32242) except for deletion of amino acids 975-986. The invention
 CC relates to a recombinant or isolated integrin heterodimer comprising the
 CC alpha10 subunit in association with subunit beta (especially beta-1). The
 CC heterodimer, subunit alpha-10 or splice variant can be used as a marker
 CC or target of all types of cells, e.g. of chondrocytes, osteoblasts and
 CC fibroblasts. They can also be used for treating pathological conditions
 CC involving ISa10, such as damage to cartilage, trauma, rheumatoid
 CC arthritis or osteoarthritis; for detecting the formation of cartilage
 CC during embryonal development, physiological or therapeutic repair of
 CC cartilage, or detecting regeneration of cartilage or chondrocytes during
 CC transplantation of cartilage or chondrocytes; for selection and analysis
 CC or for sorting, isolating or purification of chondrocytes and for in
 CC vitro studies of differentiation of chondrocytes; and as a target for
 CC anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or
 CC other tissues where adhesion impairs the function of the tissue (all
 CC claimed). ISa10 binding entities can be used to determine the
 CC differentiation-state of cells during embryonic development, angiogenesis
 CC or development of cancer, in pathological conditions such as rheumatoid
 CC arthritis, osteoarthritis or cancer, in tissue regeneration or in
 CC therapeutic and physiological repair of cartilage (claimed). A
 CC vaccine comprising the integrin heterodimer or subunit alpha-10 is also
 CC claimed. ISa10 polynucleotides, vectors, host cells and methods of
 CC producing recombinant ISa10 are also claimed
 XX
 XX Sequence 1132 AA;
 SQ
 Query Match 1.1%; Score 13; DB 3; Length 1132;
 Best Local Similarity 100.0%; Pred. No. 0.0039; Mismatches 0; Gaps 0;
 Matches 13; Conservative 0; Indels 0; Gaps 0;
 QY 166 VIVLDGNSIYPW 178
 DB 169 VIVLDGNSIYPW 181
 RESULT 73
 AAB64657
 ID AAB64657 standard; protein; 1152 AA.
 XX
 XX AAB64657;
 AC
 XX 22-MAR-2001 (first entry)
 DT
 DE Human secreted protein BLAST search protein SEQ ID NO: 167.
 DE
 XX Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
 XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200077197-A1.
 PN
 XX 21-DEC-2000.
 PD
 XX 01-JUN-2000; 2000WO-US014934.
 PF
 XX 11-JUN-1999; 99US-0138599P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 PA
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 PI
 XX WPI; 2001-032312/04.
 DR
 XX Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 PT
 XX Disclosure; Page 543-546; 589pp; English.
 PS
 XX The invention relates to the isolation of Genes AAF32757-F32803 encoding
 CC the human secreted proteins AAB64549-B64594. The sequence is a search
 CC result from a BLASTX homology search. The genes and proteins are useful
 CC for preventing, ameliorating or treating medical conditions, e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
 CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
 CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
 CC such as myocardial ischaemia; (d) wound healing; (e) neurological
 CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
 CC such as viral, bacterial, fungal and parasitic infections
 XX
 XX Sequence 1152 AA;
 SQ
 Query Match 1.1%; Score 13; DB 4; Length 1152;
 Best Local Similarity 100.0%; Pred. No. 0.004; Mismatches 0; Gaps 0;
 Matches 13; Conservative 0; Indels 0; Gaps 0;
 QY 166 VIVLDGNSIYPW 178
 DB 169 VIVLDGNSIYPW 181

CC 10 (ISa10). A splice variant is given in AAY32243. The invention relates
CC to a recombinant or isolated integrin heterodimer comprising the alpha10
CC subunit in association with subunit beta (especially beta-1). The
CC heterodimer and the subunit alpha-10 can be used as markers or targets of
CC all types of cells, e.g. of chondrocytes, osteoblasts and fibroblasts.
CC They can also be used: for treating pathological conditions involving
CC ISa10, such as damage to cartilage, trauma, rheumatoid arthritis or
CC osteoarthritis; for detecting the formation of cartilage during embryonal
CC development, physiological or therapeutic repair of cartilage, or
CC detecting regeneration of cartilage or chondrocytes during
CC transplantation of cartilage or chondrocytes; for selection and analysis
CC or for sorting, isolating or purification of chondrocytes and for in
CC vitro studies of differentiation of chondrocytes; and as a target for
CC anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or
CC other tissues where adhesion impairs the function of the tissue (all
CC claimed). ISa10 binding entities can be used to determine the
CC differentiation-state of cells during embryonic development, angiogenesis
CC or development of cancer, in pathological conditions such as rheumatoid
CC arthritis, osteoarthritis or cancer, in tissue regeneration or in
CC therapeutic and physiological repair of cartilage (claimed). A
CC vaccine comprising the integrin heterodimer or subunit alpha-10 is also
CC claimed. ISa10 polynucleotides, vectors, host cells and methods of
CC producing recombinant ISa10 are also claimed
XX
SQ Sequence 1167 AA;

Query Match 1.1%; Score 13; DB 3; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 166 VIVLDGNSIYPW 178
Db 169 VIVLDGNSIYPW 181
|||||

Search completed: September 21, 2004, 13:31:38
Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:42:08 ; Search time 79 Seconds
(without alignments)
4829.194 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRRPGLDTPKVL 1188

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 1342398 seqs, 321133274 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	926	77.9	1188	15	US-10-291-265-338
2	911	76.7	1188	15	US-10-291-265-810
3	911	76.7	1189	10	US-09-984-130-35
4	911	76.7	1189	10	US-09-836-353A-35
5	911	76.7	1189	12	US-10-262-839-6
6	810	68.2	1034	10	US-09-984-130-43
7	810	68.2	1034	10	US-09-836-353A-43
8	753	63.4	1120	12	US-10-262-839-6
9	489	41.2	707	9	US-09-764-870-313
10	489	41.2	707	14	US-10-125-540-313
11	245	20.6	437	15	US-10-108-260A-3386
12	193	16.2	193	10	US-09-805-354-8
13	193	16.2	193	12	US-09-758-493-8
14	193	16.2	193	14	US-10-144-259-8
15	88	7.4	103	9	US-09-764-870-472

16	88	7.4	103	11	US-09-764-875-1193	Sequence 1193, Ap
17	88	7.4	103	14	US-10-125-540-472	Sequence 472, App
18	42	3.5	545	42	US-10-144-259-25	Sequence 25, Appl
19	42	3.5	545	10	US-09-866-050A-500	Sequence 500, App
20	42	3.5	688	10	US-09-866-050A-624	Sequence 624, App
21	42	3.5	696	10	US-09-866-050A-501	Sequence 501, App
22	33	2.8	33	10	US-09-984-130-102	Sequence 102, App
23	33	2.8	33	10	US-09-836-353A-102	Sequence 102, App
24	21	1.8	158	10	US-09-866-050A-391	Sequence 391, App
25	17	1.4	85	10	US-09-866-050A-390	Sequence 390, App
26	15	1.3	148	9	US-09-764-870-284	Sequence 284, App
27	15	1.3	148	9	US-09-764-870-444	Sequence 444, App
28	15	1.3	148	14	US-10-125-540-284	Sequence 284, App
29	15	1.3	148	14	US-10-125-540-444	Sequence 444, App
30	15	1.3	132	12	US-10-346-863-57	Sequence 57, Appl
31	15	1.3	132	16	US-10-474-833-59	Sequence 59, Appl
32	15	1.3	132	16	US-10-474-833-60	Sequence 60, Appl
33	15	1.3	132	16	US-10-474-833-61	Sequence 61, Appl
34	15	1.3	132	16	US-09-805-354-5	Sequence 5, Appl
35	15	1.3	132	16	US-09-758-493-5	Sequence 5, Appl
36	15	1.3	132	16	US-10-144-259-5	Sequence 5, Appl
37	15	1.3	212	9	US-09-996-738-5	Sequence 5, Appl
38	15	1.3	212	9	US-09-996-738-6	Sequence 6, Appl
39	15	1.3	214	12	US-10-625-260-5	Sequence 5, Appl
40	15	1.3	214	12	US-10-625-260-6	Sequence 6, Appl
41	15	1.3	214	12	US-10-625-260-9	Sequence 9, Appl
42	15	1.3	214	13	US-10-061-658-5	Sequence 5, Appl
43	15	1.3	214	13	US-10-061-658-6	Sequence 6, Appl
44	15	1.3	214	13	US-10-061-658-9	Sequence 9, Appl
45	15	1.3	214	16	US-10-474-832-63	Sequence 63, Appl
46	15	1.3	214	16	US-10-474-832-64	Sequence 64, Appl
47	15	1.3	1151	10	US-09-984-130-103	Sequence 103, App
48	15	1.3	1151	10	US-09-836-353A-103	Sequence 103, App
49	15	1.3	1179	12	US-09-918-715-250	Sequence 250, App
50	15	1.3	1180	12	US-09-918-715-307	Sequence 307, App
51	13	1.1	195	10	US-09-805-354-7	Sequence 7, Appl
52	13	1.1	195	12	US-09-758-493-7	Sequence 7, Appl
53	13	1.1	195	14	US-10-144-259-7	Sequence 7, Appl
54	13	1.1	1167	16	US-10-741-601-531	Sequence 531, App
55	13	1.1	1177	16	US-10-741-601-532	Sequence 532, App
56	11	0.9	43	14	US-10-144-259-26	Sequence 26, Appl
57	11	0.9	1147	12	US-10-336-603A-42	Sequence 42, Appl
58	11	0.9	1181	12	US-10-211-462-187	Sequence 187, App
59	11	0.9	1181	14	US-10-160-354-2	Sequence 2, Appl
60	11	0.9	1181	15	US-10-295-027-1286	Sequence 1286, Ap
61	9	0.8	164	15	US-10-108-260A-3415	Sequence 3415, Ap
62	9	0.8	312	16	US-10-038-854-102	Sequence 102, App
63	9	0.8	312	16	US-10-369-493-2470	Sequence 2470, Ap
64	8	0.7	17	9	US-09-350-259-20	Sequence 20, Appl
65	8	0.7	17	10	US-09-891-943-20	Sequence 20, Appl
66	8	0.7	43	14	US-10-144-259-27	Sequence 27, Appl
67	8	0.7	67	9	US-09-978-295A-85	Sequence 85, Appl
68	8	0.7	67	9	US-09-978-697-85	Sequence 85, Appl
69	8	0.7	67	9	US-09-978-192A-85	Sequence 85, Appl
70	8	0.7	67	9	US-09-999-832A-85	Sequence 85, Appl
71	8	0.7	67	10	US-09-978-183-85	Sequence 85, Appl
72	8	0.7	67	10	US-09-978-608A-85	Sequence 85, Appl
73	8	0.7	67	10	US-09-978-585A-85	Sequence 85, Appl
74	8	0.7	67	10	US-09-978-191A-85	Sequence 85, Appl
75	8	0.7	67	10	US-09-978-403A-85	Sequence 85, Appl
76	8	0.7	67	10	US-09-978-564A-85	Sequence 85, Appl
77	8	0.7	67	10	US-09-999-833A-85	Sequence 85, Appl
78	8	0.7	67	10	US-09-981-915A-85	Sequence 85, Appl
79	8	0.7	67	10	US-09-978-824-85	Sequence 85, Appl
80	8	0.7	67	10	US-09-918-585A-85	Sequence 85, Appl
81	8	0.7	67	10	US-09-978-423A-85	Sequence 85, Appl
82	8	0.7	67	10	US-09-978-193A-85	Sequence 85, Appl
83	8	0.7	67	10	US-09-999-830A-85	Sequence 85, Appl
84	8	0.7	67	10	US-09-978-757A-85	Sequence 85, Appl
85	8	0.7	67	10	US-09-978-187B-85	Sequence 85, Appl
86	8	0.7	67	10	US-09-978-643A-85	Sequence 85, Appl
87	8	0.7	67	10	US-09-978-375A-85	Sequence 85, Appl
88	8	0.7	67	10	US-09-978-298A-85	Sequence 85, Appl

89	8	0.7	67	10	US-09-978-188A-85	Sequence 85, Appl
90	8	0.7	67	10	US-09-978-681A-85	Sequence 85, Appl
91	8	0.7	67	10	US-09-978-194A-85	Sequence 85, Appl
92	8	0.7	67	10	US-09-999-829A-85	Sequence 85, Appl
93	8	0.7	67	10	US-09-978-299A-85	Sequence 85, Appl
94	8	0.7	67	10	US-09-978-544A-85	Sequence 85, Appl
95	8	0.7	67	10	US-09-978-565A-85	Sequence 85, Appl
96	8	0.7	67	10	US-09-978-802A-85	Sequence 85, Appl
97	8	0.7	67	12	US-10-164-749A-85	Sequence 85, Appl
98	8	0.7	67	12	US-10-221-278-324	Sequence 324, App
99	8	0.7	67	12	US-09-999-831A-85	Sequence 85, Appl
100	8	0.7	67	12	US-10-013-917A-85	Sequence 85, Appl

ALIGNMENTS

```

RESULT 1
US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al.
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-291-265-338

```

Query Match	77.9%;	Score 926;	DB 15;	Length 1188;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1126;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps	0;			
Qy	1	MDLPRLVVAWALSALPQGTDTFENMDRKPRVPGSRTAFGTVQOCHDISGNKMLVGA	60	
Db	1	MDLPRLVVAWALSALPQGTDTFENMDRKPRVPGSRTAFGTVQOCHDISGNKMLVGA	60	
Qy	61	PLETNGYQKTGDVYKCPVIGHGCTKLNLRVTLSNVSEKXDNMRLGLSLATNPKNSFLA	120	
Db	61	PLETNGYQKTGDVYKCPVIGHGCTKLNLRVTLSNVSEKXDNMRLGLSLATNPKNSFLA	120	
Qy	121	CSPLWSHEGSSYYITGMCNRVNSNFRFSKTVAPALORCQTYMDIVIVLDGNSIYPWVE	180	
Db	121	CSPLWSHEGSSYYITGMCNRVNSNFRFSKTVAPALORCQTYMDIVIVLDGNSIYPWVE	180	
Qy	181	VQHFVFLNLIKPYIGPGQIQGVQYQGVQEDVHVEPHLNDYRSKDVVEAASHTEQRGGTET	240	
Db	181	VQHFVFLNLIKPYIGPGQIQGVQYQGVQEDVHVEPHLNDYRSKDVVEAASHTEQRGGTET	240	
Qy	241	RTAFGIEFARSEAFQKGRGKAKKMWIVITDGHSDSPDLEKVIQOQSRDNVTRYAVAVL	300	
Db	241	RTAFGIEFARSEAFQKGRGKAKKMWIVITDGHSDSPDLEKVIQOQSRDNVTRYAVAVL	300	
Qy	301	GYNRRGINPEFLNEIKYIASDDPKDHFNVNTEAALKDIDVALGRIFSLGEGTNKNET	360	
Db	301	GYNRRGINPEFLNEIKYIASDDPKDHFNVNTEAALKDIDVALGRIFSLGEGTNKNET	360	
Qy	361	SPGLEMSQTGFSSHVVEDQVLLGAVGAYDMNGAVLKETSAGKVIPIRREGLYKFEFPEELKN	420	

361	Db	SPGLEMSQTFSSHVVEDGVLGAVGAYDWNAGVLTAKETSGAKVIPURESYLKEFPBELKN	420
421	Qy	HGAYLGTYTTSVSSRQGRVYVAGAPRFNHTKGVILFTMHNRSLSITHQAMRQQQTSYF	480
421	Db	HGAYLGTYTTSVSSRQGRVYVAGAPRFNHTKGVILFTMHNRSLSITHQAMRQQQTSYF	480
481	Qy	GSEITSDVDIGDGVTDVLLVGPAPMTYNEGERGKVYVYELRQNRFYNGTLKDSHSYQNA	540
481	Db	GSEITSDVDIGDGVTDVLLVGPAPMTYNEGERGKVYVYELRQNRFYNGTLKDSHSYQNA	540
541	Qy	RGSSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFGRGSILTKPKORITASELATG	600
541	Db	RGSSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFGRGSILTKPKORITASELATG	600
601	Qy	LOYFGCSIHQQLDNLNEDGLIDLAVGALGNNAVILWSRPVVQINASLHFEPBSKINIFHRDCK	660
601	Db	LOYFGCSIHQQLDNLNEDGLIDLAVGALGNNAVILWSRPVVQINASLHFEPBSKINIFHRDCK	660
661	Qy	RGRDQATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV	720
661	Db	RGRDQATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV	720
721	Qy	LSSGGEELCERINFHVLDTADVVKPTFSVEYSLEDDPHGPMLLDGPWTTLLRVSPVFWNG	780
721	Db	LSSGGEELCERINFHVLDTADVVKPTFSVEYSLEDDPHGPMLLDGPWTTLLRVSPVFWNG	780
781	Qy	CNEDEHCVPDLVLDARSDIPTAMEYQORVLRPAQDCSAYTILSFDTTVPFIIESTORVAV	840
781	Db	CNEDEHCVPDLVLDARSDIPTAMEYQORVLRPAQDCSAYTILSFDTTVPFIIESTORVAV	840
841	Qy	EATLENRGENASTVLNISOANLOPASILQKEDSDGSIETCYNEERLRQVCNVSYPFF	900
841	Db	EATLENRGENASTVLNISOANLOPASILQKEDSDGSIETCYNEERLRQVCNVSYPFF	900
901	Qy	RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDSNERDSTKEDNVAPLRPHLYEADVLFT	960
901	Db	RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDSNERDSTKEDNVAPLRPHLYEADVLFT	960
961	Qy	RSSSIASHYEWKUNSSLERVDGIGPPFSCIFRQNLGLFPIHGMMKITTPIATRSNRL	1020
961	Db	RSSSIASHYEWKUNSSLERVDGIGPPFSCIFRQNLGLFPIHGMMKITTPIATRSNRL	1020
1021	Qy	KURDLFTDEANTSCNIWGNSTYRPTFVEEDLRAPQLNHSNSDVVSINCNTRLVFNQEI	1080
1021	Db	KURDLFTDEANTSCNIWGNSTYRPTFVEEDLRAPQLNHSNSDVVSINCNTRLVFNQEI	1080
1081	Qy	NPHLLGNLWLRSLKALKYXSMKIMVNAALQORFHSPIIFREDDPSQI	1128
1081	Db	NPHLLGNLWLRSLKALKYXSMKIMVNAALQORFHSPIIFREDDPSQI	1128

```

RESULT 2
US-10-291-265-910
; Sequence #10, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: NO. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785),
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 810
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-810

Query Match 76.7%; Score 911; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVALSLLWPGFTDTFNMTRKPRVIPSRTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRGLVAVALSLLWPGFTDTFNMTRKPRVIPSRTAFPGYTVQOHDISGNKWLAVGA 60

QY 61 PLETNQYQKTDGVYKCPVTHGCTKLNLRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120
DB 61 PLETNQYQKTDGVYKCPVTHGCTKLNLRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120

QY 121 CSPLSWHECGSSYYTTCGSRVNSNFRPSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180
DB 121 CSPLSWHECGSSYYTTCGSRVNSNFRPSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180

QY 181 VQHFLINILKFFYIGPQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
DB 181 VQHFLINILKFFYIGPQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240

QY 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPDLEKVIQCSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPDLEKVIQCSERDNVTRYAVAVL 300

QY 301 GYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGNTKNET 360
DB 301 GYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGNTKNET 360

QY 361 SFGLEMSQTFGSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPIRESYLKEPPELKN 420
DB 361 SFGLEMSQTFGSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPIRESYLKEPPELKN 420

QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTMHNRSLTTHQMRGQOIGSYF 480
DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTMHNRSLTTHQMRGQOIGSYF 480

QY 481 GSEITSVDIDGGVTDVLLVGAAPMYNEGREGKVVYVYELQNRVYNGTLDKSHSYQNA 540
DB 481 GSEITSVDIDGGVTDVLLVGAAPMYNEGREGKVVYVYELQNRVYNGTLDKSHSYQNA 540

QY 541 RFGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFHGFSGSLKTPKQRTASELATG 600
DB 541 RFGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFHGFSGSLKTPKQRTASELATG 600

QY 601 LQYFGCSIHGOLDNEDGLIDLAVGALGNVILWSRPVQINASLHFEFSKINIFHRDCK 660
DB 601 LQYFGCSIHGOLDNEDGLIDLAVGALGNVILWSRPVQINASLHFEFSKINIFHRDCK 660

QY 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
DB 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720

QY 721 LLSGGELCERINFHVLDTADYKPVTFVSVEISLEDPDHGPMLDGOWPTTLVSVFPFWNG 780
DB 721 LLSGGELCERINFHVLDTADYKPVTFVSVEISLEDPDHGPMLDGOWPTTLVSVFPFWNG 780

QY 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQDCSAYTISFDTVTFIIESTRQRAV 840
DB 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQDCSAYTISFDTVTFIIESTRQRAV 840

QY 841 EATLENRGENAYSTVLNISOSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSYPPF 900
DB 841 EATLENRGENAYSTVLNISOSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSYPPF 900

QY 901 RAKAKVAFRLD 911
DB 901 RAKAKVAFRLD 911

RESULT 3
US-09-984-130-35
; Sequence 35, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P44892
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-35

Query Match 76.7%; Score 911; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVALSLLWPGFTDTFNMTRKPRVIPSRTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRGLVAVALSLLWPGFTDTFNMTRKPRVIPSRTAFPGYTVQOHDISGNKWLAVGA 60

QY 61 PLETNQYQKTDGVYKCPVTHGCTKLNLRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120
DB 61 PLETNQYQKTDGVYKCPVTHGCTKLNLRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120

QY 121 CSPLSWHECGSSYYTTCGSRVNSNFRPSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180
DB 121 CSPLSWHECGSSYYTTCGSRVNSNFRPSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180

QY 181 VQHFLINILKFFYIGPQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
DB 181 VQHFLINILKFFYIGPQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240

QY 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPDLEKVIQCSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGHSDSPDLEKVIQCSERDNVTRYAVAVL 300

QY 301 GYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGNTKNET 360
DB 301 GYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLGNTKNET 360

QY 361 SFGLEMSQTFGSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPIRESYLKEPPELKN 420
DB 361 SFGLEMSQTFGSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPIRESYLKEPPELKN 420

QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTMHNRSLTTHQMRGQOIGSYF 480
DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTMHNRSLTTHQMRGQOIGSYF 480

QY 481 GSEITSVDIDGGVTDVLLVGAAPMYNEGREGKVVYVYELQNRVYNGTLDKSHSYQNA 540
DB 481 GSEITSVDIDGGVTDVLLVGAAPMYNEGREGKVVYVYELQNRVYNGTLDKSHSYQNA 540

QY 541 RFGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFHGFSGSLKTPKQRTASELATG 600
DB 541 RFGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFHGFSGSLKTPKQRTASELATG 600

```
QY 601 LQVFGCSIHGQDLNEDGLIDLAVALGNVILWSPVQINASLHFEPKINIFHRDCK 660
Db 601 LQVFGCSIHGQDLNEDGLIDLAVALGNVILWSPVQINASLHFEPKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRTPRAHLDEGGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRTPRAHLDEGGDRFTNRAV 720
QY 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLEDDPGMPMDGWPPTTLRVSVYFPWNG 780
Db 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLEDDPGMPMDGWPPTTLRVSVYFPWNG 780
QY 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIIBSTRQAV 840
Db 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIIBSTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911
```

RESULT 4

```
US-09-836-353A-35
; Sequence 35, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-35
```

```
Query Match 76.7%; Score 911; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDPLRGLVAVWALSMLPGFTDFNWDTRKPRVIPSRTAFPGYTVQOHDISGNKWLTVGA 60
Db 1 MDPLRGLVAVWALSMLPGFTDFNWDTRKPRVIPSRTAFPGYTVQOHDISGNKWLTVGA 60
QY 61 PLETNQYQKTDGVYKCPVIGHNCTKLNLRVTLNVSERKONMRLGLSLATNPKDNSFLA 120
Db 61 PLETNQYQKTDGVYKCPVIGHNCTKLNLRVTLNVSERKONMRLGLSLATNPKDNSFLA 120
QY 121 CSPLSWHEGCSYITTCMSRVNSFRSKTVAPALQRCQTMIDIVIVLDGNSIYPPWE 180
Db 121 CSPLSWHEGCSYITTCMSRVNSFRSKTVAPALQRCQTMIDIVIVLDGNSIYPPWE 180
QY 181 VQHFLINILKXYIGPGQIQVGVQYGEDVHFEHLNDRYSVKDVVEAASHIEQGGTET 240
Db 181 VQHFLINILKXYIGPGQIQVGVQYGEDVHFEHLNDRYSVKDVVEAASHIEQGGTET 240
QY 241 RTAFGIEFARSAFQKGGKGAKKVMIVITDGEHSDSPLEKVIQCSERDWNTRYAVAVL 300
Db 241 RTAFGIEFARSAFQKGGKGAKKVMIVITDGEHSDSPLEKVIQCSERDWNTRYAVAVL 300
```

```
QY 301 GYNRRGINPETELNEIKYIASDPDDKHFNFTDEAALKDIDVALGDRIFSLEGTNKNET 360
Db 301 GYNRRGINPETELNEIKYIASDPDDKHFNFTDEAALKDIDVALGDRIFSLEGTNKNET 360
QY 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPPELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPPELKN 420
QY 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQAMRGGQIGSYF 480
Db 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQAMRGGQIGSYF 480
QY 481 GSIITSVIDIDGDTVLLVGAAPMYNEGBERGKVVYELRQNRFYVNGTLKDSHSYQNA 540
Db 481 GSIITSVIDIDGDTVLLVGAAPMYNEGBERGKVVYELRQNRFYVNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQRTASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKQRTASELATG 600
QY 601 LQVFGCSIHGQDLNEDGLIDLAVALGNVILWSPVQINASLHFEPKINIFHRDCK 660
Db 601 LQVFGCSIHGQDLNEDGLIDLAVALGNVILWSPVQINASLHFEPKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRTPRAHLDEGGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRTPRAHLDEGGDRFTNRAV 720
QY 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLEDDPGMPMDGWPPTTLRVSVYFPWNG 780
Db 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLEDDPGMPMDGWPPTTLRVSVYFPWNG 780
QY 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIIBSTRQAV 840
Db 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIIBSTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911
```

RESULT 5

```
US-10-262-839-4
; Sequence 4, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
```



```
QY 121 CSPLWSEHCSSYYTTCMCSRNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 121 CSPLWSEHCSSYYTTCMCSRNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
QY 181 VQHFLINILKFFYIGPQIQGVVQVQGVGDEHFFHNDYRSKDVVAAASHIEQRGGET 240
Db 181 VQHFLINILKFFYIGPQIQGVVQVQGVGDEHFFHNDYRSKDVVAAASHIEQRGGET 240
QY 241 RTAFGIEFARSAFQGGKGGKAKKMWITTDGSHSDSPLEKVIQCSERDNTVAVAVL 300
Db 241 RTAFGIEFARSAFQGGKGGKAKKMWITTDGSHSDSPLEKVIQCSERDNTVAVAVL 300
QY 301 GYNNRGINPETFLNEIKVIASDPDDKHFNFNTDEAALKDIIVDALGDRIFSLGNTKNET 360
Db 301 GYNNRGINPETFLNEIKVIASDPDDKHFNFNTDEAALKDIIVDALGDRIFSLGNTKNET 360
QY 361 SFGLMSQTFSSHVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEPPEELKN 420
Db 361 SFGLMSQTFSSHVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEPPEELKN 420
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSILTIHQAMEGQIGSYF 480
Db 421 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSILTIHQAMEGQIGSYF 480
QY 481 GSEITSVDIDGCVTDVLLVAGAPMYFNEGRGKVVYVELRQNRVYNGTLKDSHYQNA 540
Db 481 GSEITSVDIDGCVTDVLLVAGAPMYFNEGRGKVVYVELRQNRVYNGTLKDSHYQNA 540
QY 541 RFSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFGRSILTKPKRITASELATG 600
Db 541 RFSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFGRSILTKPKRITASELATG 600
QY 601 LQYFGCSIHGQDLNEDGLIDLAGALGNVILWSRPVQINASLHFPFSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAGALGNVILWSRPVQINASLHFPFSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
QY 721 LSSGQELCERINFHVLDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVPPFWNG 780
Db 721 LSSGQELCERINFHVLDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVPPFWNG 780
QY 781 CNEDEHCVPDLVLDARSDLPTAMEYQCVLRKPAQDCSAYTLSFDTTTFIIESTRQAV 840
Db 781 CNEDEHCVPDLVLDARSDLPTAMEYQCVLRKPAQDCSAYTLSFDTTTFIIESTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSIETCNEERLRKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSIETCNEERLRKQVCNVSYPFF 900
QY 901 RAKAKVAFRILD 911
Db 901 RAKAKVAFRILD 911
```

RESULT 7

```
US-09-836-353A-43
; Sequence 43, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P/489P1
; CURRENT APPLICATION NUMBER: US/09/836, 353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
```

```
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-43
```

```
Query Match 68.2%; Score 810; DB 10; Length 1034;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSMPGFTDTFNMDTRPRVTPGSRATFAFGYTVQOHDISGNKMLVGA 60
Db 1 MDLPRGLVAVWALSMPGFTDTFNMDTRPRVTPGSRATFAFGYTVQOHDISGNKMLVGA 60
QY 61 PLETNGYQKTGDYVKCPVIHGNCTKLNLRVTLSNYSERKDNRLGLSLATNPKNLSFLA 120
Db 61 PLETNGYQKTGDYVKCPVIHGNCTKLNLRVTLSNYSERKDNRLGLSLATNPKNLSFLA 120
QY 121 CSPLWSEHCSSYYTTCMCSRNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 121 CSPLWSEHCSSYYTTCMCSRNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
QY 181 VQHFLINILKFFYIGPQIQGVVQVQGVGDEHFFHNDYRSKDVVAAASHIEQRGGET 240
Db 181 VQHFLINILKFFYIGPQIQGVVQVQGVGDEHFFHNDYRSKDVVAAASHIEQRGGET 240
QY 241 RTAFGIEFARSAFQGGKGGKAKKMWITTDGSHSDSPLEKVIQCSERDNTVAVAVL 300
Db 241 RTAFGIEFARSAFQGGKGGKAKKMWITTDGSHSDSPLEKVIQCSERDNTVAVAVL 300
QY 301 GYNNRGINPETFLNEIKVIASDPDDKHFNFNTDEAALKDIIVDALGDRIFSLGNTKNET 360
Db 301 GYNNRGINPETFLNEIKVIASDPDDKHFNFNTDEAALKDIIVDALGDRIFSLGNTKNET 360
QY 361 SFGLMSQTFSSHVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEPPEELKN 420
Db 361 SFGLMSQTFSSHVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEPPEELKN 420
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSILTIHQAMEGQIGSYF 480
Db 421 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSILTIHQAMEGQIGSYF 480
QY 481 GSEITSVDIDGCVTDVLLVAGAPMYFNEGRGKVVYVELRQNRVYNGTLKDSHYQNA 540
Db 481 GSEITSVDIDGCVTDVLLVAGAPMYFNEGRGKVVYVELRQNRVYNGTLKDSHYQNA 540
QY 541 RFSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFGRSILTKPKRITASELATG 600
Db 541 RFSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFGRSILTKPKRITASELATG 600
QY 601 LQYFGCSIHGQDLNEDGLIDLAGALGNVILWSRPVQINASLHFPFSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAGALGNVILWSRPVQINASLHFPFSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
QY 721 LSSGQELCERINFHVLDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVPPFWNG 780
Db 721 LSSGQELCERINFHVLDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVPPFWNG 780
QY 781 CNEDEHCVPDLVLDARSDLPTAMEYQCVLRKPAQDCSAYTLSFDTTTFIIESTRQAV 840
Db 781 CNEDEHCVPDLVLDARSDLPTAMEYQCVLRKPAQDCSAYTLSFDTTTFIIESTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSIETCNEERLRKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSIETCNEERLRKQVCNVSYPFF 900
QY 901 RAKAKVAFRILD 911
```


Db 901 RAKAKVAFELD 911

|||||

RESULT 8

US-10-262-839-6

; Sequence 6, Application US/10262839

; Publication No. US20040039877A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook, John,

; APPLICANT: Anderson, David W.,

; APPLICANT: Boldog, Ferenc,

; APPLICANT: Burgess, Catherine,

; APPLICANT: Catterton, Elina,

; APPLICANT: Edinger, Shlomit,

; APPLICANT: Ellerman, Karen,

; APPLICANT: Gerlach, Valerie,

; APPLICANT: Gorman, Linda,

; APPLICANT: Guo, Xiaojia,

; APPLICANT: Ji, Weizhen,

; APPLICANT: Kekuda, Ramesh,

; APPLICANT: Leach, Martin,

; APPLICANT: Li, Li,

; APPLICANT: Miller, Charles,

; APPLICANT: Patturajan, Meera,

; APPLICANT: Reiger, Daniel,

; APPLICANT: Rotenberg, Mark,

; APPLICANT: Shimkets, Richard,

; APPLICANT: Smithson, Glennda,

; APPLICANT: Spytek, Kimberly,

; APPLICANT: Taupier, Corinne, jr.,

; APPLICANT: Vernet, Corine,

; APPLICANT: Voss, Edward,

; APPLICANT: Zerhusen, Brian,

; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-462A

; CURRENT APPLICATION NUMBER: US/10/262,839

; CURRENT FILING DATE: 2002-10-01

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,101

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/371,972

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/327,342

; PRIOR FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 60/328,044

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,849

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/374,738

; PRIOR FILING DATE: 2002-04-23

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 367

; SOFTWARE: Curaseqlust version 0.1

; SEQ ID NO 6

; LENGTH: 1120

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-262-839-6

Query Match 63.4%; Score 753; DB 12; Length 1120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQYMDIVIVLDSGNSIYPWVEVQHFLINILKKFYLGPGQIQGVVQYGEDVVFHFLND 218

Db 90 CQYMDIVIVLDSGNSIYPWVEVQHFLINILKKFYLGPGQIQGVVQYGEDVVFHFLND 149

QY 219 YRSVKDVEAASHIEQRCGTETRTAFGIEFARSEAFQKGRGAKKVMIVITDGEHSDP 278

Db 150 YRSVKDVEAASHIEQRCGTETRTAFGIEFARSEAFQKGRGAKKVMIVITDGEHSDP 209

QY 279 DLEKVIQOSERDNNVYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVDTAAL 338

Db 210 DLEKVIQOSERDNNVYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVDTAAL 269

QY 339 KDIVDALGDRIFSLEGTKNNETSFGLEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 398

Db 270 KDIVDALGDRIFSLEGTKNNETSFGLEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 329

QY 399 SAGKVIPLRESYLKEPPEELKNHGAIVGTVTSVSSRQGRVYVACAPFNHTGKILFT 458

Db 330 SAGKVIPLRESYLKEPPEELKNHGAIVGTVTSVSSRQGRVYVACAPFNHTGKILFT 389

QY 459 MHNRSLSLTHQAMRGQOIGSYFGSEITSVDIDGCVTDVLLVGAPMYFNEGRERGVVY 518

Db 390 MHNRSLSLTHQAMRGQOIGSYFGSEITSVDIDGCVTDVLLVGAPMYFNEGRERGVVY 449

QY 519 ELQRNPFVNGTLKDSHSYQNAFSGSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFH 578

Db 450 ELQRNPFVNGTLKDSHSYQNAFSGSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFH 509

QY 579 GFSGSILKTPKQRTASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNAILSRPV 638

Db 510 GFSGSILKTPKQRTASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNAILSRPV 569

QY 639 VQINASLHFPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMD 698

Db 570 VQINASLHFPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMD 629

QY 699 ERYTPRAHLDEGGDRFTNRAVLLSSGQELCERINPHVLDTADYVKEPVSYSLEDDP 758

Db 630 ERYTPRAHLDEGGDRFTNRAVLLSSGQELCERINPHVLDTADYVKEPVSYSLEDDP 689

QY 759 HGPMLDDGWPTTLRVSVFPWNGCNEDEHCVFDLVLDARSDLPTAMEYQCRVLRKPAQDCS 818

Db 690 HGPMLDDGWPTTLRVSVFPWNGCNEDEHCVFDLVLDARSDLPTAMEYQCRVLRKPAQDCS 749

QY 819 AYTLSPDTTVPFIESTRQRVAVATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 878

Db 750 AYTLSPDTTVPFIESTRQRVAVATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 809

QY 879 IECVNEERLQKQVCNVSYPPFFRAKAKVAFRLD 911

Db 810 IECVNEERLQKQVCNVSYPPFFRAKAKVAFRLD 842

RESULT 9

US-09-764-870-313

; Sequence 313, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ14

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 313

; LENGTH: 707

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-870-313

Query Match 41.2%; Score 489; DB 9; Length 707;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MDLPRLGLVAVWALSILMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 60
DB 17 MDLPRLGLVAVWALSILMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 76
QY 61 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 120
DB 77 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 136
QY 121 CSPLWSEHCSSSYTTGMSRVNSNFRFSTKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
DB 137 CSPLWSEHCSSSYTTGMSRVNSNFRFSTKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
QY 181 VQFLNINILKKFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
DB 197 VQFLNINILKKFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 256
QY 241 RTAFGIEFARSEAFQKGRKGAKVMIVITDGSHSDSPLEKVIQOSERDNVTRYAVAVL 300
DB 257 RTAFGIEFARSEAFQKGRKGAKVMIVITDGSHSDSPLEKVIQOSERDNVTRYAVAVL 316
QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLGNTKNET 360
DB 317 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLGNTKNET 376
QY 361 SFGLEMSQTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420
DB 377 SFGLEMSQTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 436
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSILTIHQMRGQQIGSYF 480
DB 437 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSILTIHQMRGQQIGSYF 496
QY 481 GSEITSDVDIGDGVTDVLLVAGAPMYNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 540
DB 497 GSEITSDVDIGDGVTDVLLVAGAPMYNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 556
QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
DB 557 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 616
QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQINASHLHFEPSKINIFHRDCK 660
DB 617 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQINASHLHFEPSKINIFHRDCK 676
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 691
DB 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 707

```

RESULT 10

US-10-125-540-313
 ; Sequence 313, Application US/10125540
 ; Publication No. US20030059875A1
 ; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PTZ14C1
 ; CURRENT APPLICATION NUMBER: US/10/125,540
 ; CURRENT FILING DATE: 2002-04-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 646
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 313
 ; LENGTH: 707
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-125-540-313

Query Match 41.2%; Score 489; DB 14; Length 707;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MDLPRLGLVAVWALSILMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 60
DB 17 MDLPRLGLVAVWALSILMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 76
QY 61 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 120
DB 77 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 136
QY 121 CSPLWSEHCSSSYTTGMSRVNSNFRFSTKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
DB 137 CSPLWSEHCSSSYTTGMSRVNSNFRFSTKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
QY 181 VQFLNINILKKFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
DB 197 VQFLNINILKKFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 256
QY 241 RTAFGIEFARSEAFQKGRKGAKVMIVITDGSHSDSPLEKVIQOSERDNVTRYAVAVL 300
DB 257 RTAFGIEFARSEAFQKGRKGAKVMIVITDGSHSDSPLEKVIQOSERDNVTRYAVAVL 316
QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLGNTKNET 360
DB 317 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLGNTKNET 376
QY 361 SFGLEMSQTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420
DB 377 SFGLEMSQTGFSSHVVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 436
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSILTIHQMRGQQIGSYF 480
DB 437 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVILFTMHNRSILTIHQMRGQQIGSYF 496
QY 481 GSEITSDVDIGDGVTDVLLVAGAPMYNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 540
DB 497 GSEITSDVDIGDGVTDVLLVAGAPMYNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 556
QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
DB 557 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 616
QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQINASHLHFEPSKINIFHRDCK 660
DB 617 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQINASHLHFEPSKINIFHRDCK 676
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 691
DB 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 707

```

RESULT 11

US-10-108-260A-3386
 ; Sequence 3386, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No US20040005560A1el full length cdna
 ; FILE REFERENCE: H1-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3386
 ; LENGTH: 437
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-108-260A-3386

Query Match 20.6%; Score 245; DB 15; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1e-224;
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 175 IYPWVEVQHFLNINILKKFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQ 234

```
Db 175 IYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLNDYRSVKDVEAASHIEQ 234
Qy 235 RGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDPDLEKVIQOSERDNTVR 294
Db 235 RGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDPDLEKVIQOSERDNTVR 294
Qy 295 YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIVDALGDRIFSLG 354
Db 295 YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIVDALGDRIFSLG 354
Qy 355 TNKNETSFGLMSQTFSSHVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEF 414
Db 355 TNKNETSFGLMSQTFSSHVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEF 414
Qy 415 PEELK 419
Db 415 PEELK 419

RESULT 12
US-09-805-354-8
; Sequence 8, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001 US/09/805,354
; CURRENT APPLICATION NUMBER: US/09/805,354
; PRIOR FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-8

Query Match 16.2%; Score 193; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 60
Qy 219 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 278
Db 61 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 120
Qy 279 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
Db 121 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 180
Qy 339 KDIVDALGDRIFS 351
Db 181 KDIVDALGDRIFS 193

RESULT 14
US-10-144-259-8
; Sequence 8, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-8

Query Match 16.2%; Score 193; DB 14; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 60
Qy 219 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 278
Db 61 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 120
Qy 279 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
```

```
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-8

Query Match 16.2%; Score 193; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 60
Qy 219 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 278
Db 61 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 120
Qy 279 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
Db 121 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 180
Qy 339 KDIVDALGDRIFS 351
Db 181 KDIVDALGDRIFS 193

RESULT 14
US-10-144-259-8
; Sequence 8, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-8

Query Match 16.2%; Score 193; DB 14; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGGQIQGVVQGVGVHFLND 60
Qy 219 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 278
Db 61 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGESHDP 120
Qy 279 DLEKVIQOSERDNTVR YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
```

Db 121 LDKVIOQSDNTRYAVAVLGYNRRGINPFTFLNEIKYIASDDDKHFFNVDEAAL 180
Qy 339 KDIVDALGDRIFS 351
Db 181 KDIVDALGDRIFS 193

RESULT 15
US-09-764-870-472
; Sequence 472, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 472
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-472

Query Match 7.4%; Score 88; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 218
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 67

Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 16
US-09-764-875-1193
; Sequence 1193, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PUZ02
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1193
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-1193

Query Match 7.4%; Score 88; DB 11; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 218

Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 67
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 17
US-10-125-540-472
; Sequence 472, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 472
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-472

Query Match 7.4%; Score 88; DB 14; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 218
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFFHLND 67

Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 18
US-10-144-259-25
; Sequence 25, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-25

Query Match 3.5%; Score 42; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 9.2e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 335 EALKKDIVDALGDRIFSLSGTGNKNTSPGLEMSQTGFSSHV 376

```
Db 1 EAALKDIVDAGDRIFSLGKTNKNETSPGLEMSQTGFSSHVV 42

RESULT 19
US-09-866-050A-500
; Sequence 500, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-500

Query Match 3.5%; Score 42; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 9.1e-31;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 123 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 164
|||||

RESULT 20
US-09-866-050A-624
; Sequence 624, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 624
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-624

Query Match 3.5%; Score 42; DB 10; Length 688;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307
|||||

RESULT 21
US-09-866-050A-501
; Sequence 501, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-501

Query Match 3.5%; Score 42; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307
|||||

RESULT 22
US-09-984-130-102
; Sequence 102, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-102

Query Match 2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDVYKCPVHGNCCTKLNGLGRVTLNSV 96
|||||
Db 1 TNGYQKTGDVYKCPVHGNCCTKLNGLGRVTLNSV 33
|||||

RESULT 23
US-09-836-353A-102
; Sequence 102, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
```

```
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-102

Query Match      2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQXTGDKYKCPVHGNCTKLNLCGRVTLNV 96
Db 1 TNGYQXTGDKYKCPVHGNCTKLNLCGRVTLNV 33

RESULT 24
US-09-866-050A-391
; Sequence 391, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-391

Query Match      1.8%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHPEPSKINIFH 656
Db 57 RPVVQINASLHPEPSKINIFH 77

RESULT 25
US-09-866-050A-390
; Sequence 390, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U

; FILE REFERENCE: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-390

Query Match      1.4%; Score 17; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNARFGS 544
Db 28 NGTLKDSHSYQNARFGS 44

RESULT 26
US-09-764-870-284
; Sequence 284, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 284
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-284

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDGNSIYPW 178
Db 47 DIVVLGDGNSIYPW 61

RESULT 27
US-09-764-870-444
; Sequence 444, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 444
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-444

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDGNSIYPW 178
Db 47 DIVVLGDGNSIYPW 61
```

```

RESULT 28
US-10-125-540-284
; Sequence 284, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; PRIOR FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 284
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-540-284

Query Match      1.3%; Score 15; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      47 DIVVLGSGNSIYPW 61

RESULT 29
US-10-125-540-444
; Sequence 444, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; PRIOR FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 444
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-540-444

Query Match      1.3%; Score 15; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      47 DIVVLGSGNSIYPW 61

RESULT 30
US-10-346-863-57
; Sequence 57, Application US/10346863
; Publication No. US20040038325A1
; GENERAL INFORMATION:
; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN
; APPLICANT: FAGAN, RICHARD JOSEPH
; APPLICANT: GUTTERIDGE, ALEX
; TITLE OF INVENTION: ADHESION MOLECULES
; FILE REFERENCE: 674575-2001
; CURRENT APPLICATION NUMBER: US/10/346,863
; CURRENT FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: PCT/GE01/03318
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: GB 0018126.3

```

```

; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: GB 0025447.4
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-346-863-57

Query Match      1.3%; Score 15; DB 12; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      5 DIVVLGSGNSIYPW 19

RESULT 31
US-10-474-832-59
; Sequence 59, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Rat/human chimeric
; OTHER INFORMATION: I domain construct
US-10-474-832-59

Query Match      1.3%; Score 15; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      4 DIVVLGSGNSIYPW 18

RESULT 32
US-10-474-832-60
; Sequence 60, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 192
; TYPE: PRT

```

```

; ORGANISM: Rattus sp.
US-10-474-832-60

Query Match      1.3%; Score 15; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 18

RESULT 33
US-10-474-832-61
; Sequence 61, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-61

Query Match      1.3%; Score 15; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 18

RESULT 34
US-09-805-354-5
; Sequence 5, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001
; CURRENT APPLICATION NUMBER: US/09/805,354
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-5

Query Match      1.3%; Score 15; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 18

US-09-805-354-5
; Sequence 5, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-5

Query Match      1.3%; Score 15; DB 14; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      6 DIVVLGDSNSIYPW 20

RESULT 35
US-09-758-493-5
; Sequence 5, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-5

Query Match      1.3%; Score 15; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      6 DIVVLGDSNSIYPW 20

RESULT 36
US-10-144-259-5
; Sequence 5, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-5

Query Match      1.3%; Score 15; DB 14; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      6 DIVVLGDSNSIYPW 20

RESULT 37
US-09-996-738-5
; Sequence 5, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:

```



```
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougereolles, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: rat
US-09-996-738-5

Query Match      1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      22 DIVVLGSGNSIYPW 36

RESULT 38
US-09-996-738-6
; Sequence 6, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougereolles, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-996-738-6

Query Match      1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      22 DIVVLGSGNSIYPW 36

RESULT 39
US-10-625-260-5
; Sequence 5, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-625-260-5

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      22 DIVVLGSGNSIYPW 36

RESULT 40
US-10-625-260-6
; Sequence 6, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-6

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYPW 178
Db      22 DIVVLGSGNSIYPW 36

RESULT 41
US-10-625-260-9
; Sequence 9, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
```

```

; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-9

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 42
US-10-061-658-5
; Sequence 5, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-061-658-5

Query Match      1.3%; Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 43
US-10-061-658-6
; Sequence 6, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-474-832-63

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 44
US-10-061-658-9
; Sequence 9, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

Query Match      1.3%; Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 45
US-10-474-832-63
; Sequence 63, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-474-832-63

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178

```

```
Db      22 DIVVLGSGNSIYFW 36
|||||
RESULT 46
US-10-474-832-64
; Sequence 64, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-64

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      22 DIVVLGSGNSIYFW 36
|||||

RESULT 47
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      144 DIVVLGSGNSIYFW 158
|||||

RESULT 48
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      144 DIVVLGSGNSIYFW 158
|||||

RESULT 49
US-09-918-715-250
; Sequence 250, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Best Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-250

Query Match      1.3%; Score 15; DB 12; Length 1179;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      172 DIVVLGSGNSIYFW 186
|||||

RESULT 50
US-09-918-715-307
; Sequence 307, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
```

```
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107,00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 307
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Rat
US-09-918-715-307
```

```
Query Match 1.1%; Score 15; DB 12; Length 1180;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGDSNSIYPW 178
Db 172 DIVVLGDSNSIYPW 186
|||||
```

RESULT 51

```
US-09-805-354-7
; Sequence 7, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001
; CURRENT APPLICATION NUMBER: US/09/805,354
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-7
```

```
Query Match 1.1%; Score 13; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 VIVLDGNSIYPW 178
Db 8 VIVLDGNSIYPW 20
|||||
```

RESULT 52

```
US-09-758-493-7
; Sequence 7, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
```

```
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-7
```

```
Query Match 1.1%; Score 13; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 VIVLDGNSIYPW 178
Db 8 VIVLDGNSIYPW 20
|||||
```

RESULT 53

```
US-10-144-259-7
; Sequence 7, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-7
```

```
Query Match 1.1%; Score 13; DB 14; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 VIVLDGNSIYPW 178
Db 8 VIVLDGNSIYPW 20
|||||
```

RESULT 54

```
US-10-741-601-531
; Sequence 531, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-531
```

```

Query Match      1.13; Score 13; DB 16; Length 1167;
Best Local Similarity 100.0%; Pred.No. 0.0089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      166 VIVLDGNSNIYPW 178
      |||||
DB      169 VIVLDGNSNIYPW 181

RESULT 55
US-10-741-601-532
; Sequence 532, Application US/10741601
; Publication No. US20040165519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 532
; LENGTH: 1177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-532

```

```

Query Match      1.13; Score 13; DB 16; Length 1177;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      166 VIVLDGNSIYPW 178
      |||||
Db       169 VIVLDGNSIYPW 181

RESULT 56
US-10-144-259-26
; Sequence 26, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/759,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 43
; TYPE: PRT

```

```
US-10-144-259-26
Query Match          0.9%; Score 11; DB 14; Length 43;
Best Local Similarity 100.0%; Pred.No. 0.038;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      340 DIVDALGDRIF 350
        |||||
Db       6 DIVDALGDRIF 15

RESULT 57
US-10-336-603A-42
; Sequence 42; Application US/10336603A
```

```

; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Pardo, Jorge
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Alpha 2 Integrin: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-001110US
; CURRENT APPLICATION NUMBER: US/10/160,354
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/296,819
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: wild type alpha 2 integrin, CD49b
US-10-160-354-2

```

```

Query Match          0.9%; Score 11; DB 14; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516

```

```

RESULT 60
US-10-295-027-1286
; Sequence 1286, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 1286
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1286

Query Match          0.9%; Score 11; DB 15; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516

```

```

RESULT 61
US-10-108-260A-3415
; Sequence 3415, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: HL-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3415
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3415

```

```

Query Match          0.8%; Score 9; DB 15; Length 164;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLVL 1161
Db 99 LLLALLLVL 107

```

```

RESULT 62
US-10-036-854-102
; Sequence 102, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle

```

```
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Bos taurus
; US-10-038-854-102

Query Match      0.8%; Score 9; DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1154 LLLALLVLA 1162
      |||||
Db      3 LLLALLVLA 11

RESULT 63
US-10-369-493-2470
; Sequence 2470, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2470
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; US-10-369-493-2470

Query Match      0.8%; Score 9; DB 15; Length 371;
Best Local Similarity 100.0%; Pred. No. 21;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      378 DGVLLGAVG 386
      |||||
```

```
Db      69 DGVLLGAVG 77

RESULT 64
US-09-350-259-20
; Sequence 20, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
; US-09-350-259-20

Query Match      0.7%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 12;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      56 LVVGAPLE 63
      |||||
Db      1 LVVGAPLE 8

RESULT 65
US-09-891-943-20
; Sequence 20, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
; US-09-891-943-20

Query Match      0.7%; Score 8; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 12;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      56 LVVGAPLE 63
```


; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083332
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 68

US-09-978-697-85
; Sequence 85, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin J.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Klavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03

PRIOR FILING DATE: 1998-04-15	
PRIOR APPLICATION NUMBER: 60/081819	
PRIOR FILING DATE: 1998-04-15	
PRIOR APPLICATION NUMBER: 60/081952	
PRIOR FILING DATE: 1998-04-15	
PRIOR APPLICATION NUMBER: 60/081938	
PRIOR FILING DATE: 1998-04-15	
PRIOR APPLICATION NUMBER: 60/082568	
PRIOR FILING DATE: 1998-04-21	
PRIOR APPLICATION NUMBER: 60/082569	
PRIOR FILING DATE: 1998-04-21	
PRIOR APPLICATION NUMBER: 60/082704	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082804	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082700	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082797	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082796	
PRIOR FILING DATE: 1998-04-23	
PRIOR APPLICATION NUMBER: 60/083336	
PRIOR FILING DATE: 1998-04-27	
PRIOR APPLICATION NUMBER: 60/083322	
PRIOR FILING DATE: 1998-04-28	
PRIOR APPLICATION NUMBER: 60/083392	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083495	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083496	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083499	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083545	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083554	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083558	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083559	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083500	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083742	
PRIOR FILING DATE: 1998-04-30	
PRIOR APPLICATION NUMBER: 60/084366	
PRIOR FILING DATE: 1998-05-05	
PRIOR APPLICATION NUMBER: 60/084414	
PRIOR FILING DATE: 1998-05-06	
PRIOR APPLICATION NUMBER: 60/084441	
PRIOR FILING DATE: 1998-05-06	
PRIOR APPLICATION NUMBER: 60/084637	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084639	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084640	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084598	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084600	
PRIOR FILING DATE: 1998-5-07	
PRIOR APPLICATION NUMBER: 60/084627	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084643	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/085339	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085338	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085323	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085582	
PRIOR FILING DATE: 1998-05-15	

/ PRIOR APPLICATION NUMBER: 60/085700
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085689
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085579
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085580
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085573
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085704
 / PRIOR FILING DATE: 1998-05-15
 / PRIOR APPLICATION NUMBER: 60/085597

Query Match 0.7%; Score 8; DB 9; Length 67;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY / 1152 GLLLLALL 1159

Db 12 GLLLLALL 19

RESULT 69

US-09-978-192A-85
 / Sequence 85, Application US/09978192A

Patent No. US20020177553A1

GENERAL INFORMATION:

/ APPLICANT: Ashkenazi, Avi
 / APPLICANT: Baker Kevin P.
 / APPLICANT: Botstein, David
 / APPLICANT: Desnovers, Luc
 / APPLICANT: Eaton, Dan
 / APPLICANT: Ferrara, Napoleon
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Fong, Sherman
 / APPLICANT: Gac, Wei-Qiang
 / APPLICANT: Gerber, Hanspeter
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Grimaldi, J. Christopher
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Hillan, Kenneth J.
 / APPLICANT: Kljavin, Ivar J.
 / APPLICANT: Kuo, Sophia S.
 / APPLICANT: Napier, Mary A.
 / APPLICANT: Pan, James;
 / APPLICANT: Paoni, Nicholas F.
 / APPLICANT: Roy, Margaret Ann
 / APPLICANT: Shelton, David L.
 / APPLICANT: Stewart, Timothy A.
 / APPLICANT: Tumas, Daniel
 / APPLICANT: Williams, P. Mickey
 / APPLICANT: Wood, William I.
 / TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 / TITLE OF INVENTION: Acids Encoding the Same
 / FILE REFERENCE: P2630P1C9
 / CURRENT APPLICATION NUMBER: US/09/978,192A
 / CURRENT FILING DATE: 2001-10-15
 / PRIOR APPLICATION NUMBER: 09/918585
 / PRIOR FILING DATE: 2001-07-30
 / PRIOR APPLICATION NUMBER: 60/062250
 / PRIOR FILING DATE: 1997-10-17
 / PRIOR APPLICATION NUMBER: 60/064249
 / PRIOR FILING DATE: 1997-11-03
 / PRIOR APPLICATION NUMBER: 60/065311
 / PRIOR FILING DATE: 1997-11-13
 / PRIOR APPLICATION NUMBER: 60/066364
 / PRIOR FILING DATE: 1997-11-21
 / PRIOR APPLICATION NUMBER: 60/077450
 / PRIOR FILING DATE: 1998-03-10
 / PRIOR APPLICATION NUMBER: 60/077632

/ PRIOR FILING DATE: 1998-03-11
 / PRIOR APPLICATION NUMBER: 60/077641
 / PRIOR FILING DATE: 1998-03-11
 / PRIOR APPLICATION NUMBER: 60/077649
 / PRIOR FILING DATE: 1998-03-11
 / PRIOR APPLICATION NUMBER: 60/077791
 / PRIOR FILING DATE: 1998-03-12
 / PRIOR APPLICATION NUMBER: 60/078004
 / PRIOR FILING DATE: 1998-03-13
 / PRIOR APPLICATION NUMBER: 60/078886
 / PRIOR FILING DATE: 1998-03-20
 / PRIOR APPLICATION NUMBER: 60/078936
 / PRIOR FILING DATE: 1998-03-20
 / PRIOR APPLICATION NUMBER: 60/078910
 / PRIOR FILING DATE: 1998-03-20
 / PRIOR APPLICATION NUMBER: 60/078939
 / PRIOR FILING DATE: 1998-03-20
 / PRIOR APPLICATION NUMBER: 60/079294
 / PRIOR FILING DATE: 1998-03-25
 / PRIOR APPLICATION NUMBER: 60/079656
 / PRIOR FILING DATE: 1998-03-26
 / PRIOR APPLICATION NUMBER: 60/079664
 / PRIOR FILING DATE: 1998-03-27
 / PRIOR APPLICATION NUMBER: 60/079689
 / PRIOR FILING DATE: 1998-03-27
 / PRIOR APPLICATION NUMBER: 60/079663
 / PRIOR FILING DATE: 1998-03-27
 / PRIOR APPLICATION NUMBER: 60/079728
 / PRIOR FILING DATE: 1998-03-27
 / PRIOR APPLICATION NUMBER: 60/079786
 / PRIOR FILING DATE: 1998-03-27
 / PRIOR APPLICATION NUMBER: 60/079920
 / PRIOR FILING DATE: 1998-03-30
 / PRIOR APPLICATION NUMBER: 60/079923
 / PRIOR FILING DATE: 1998-03-30
 / PRIOR APPLICATION NUMBER: 60/080105
 / PRIOR FILING DATE: 1998-03-31
 / PRIOR APPLICATION NUMBER: 60/080107
 / PRIOR FILING DATE: 1998-03-31
 / PRIOR APPLICATION NUMBER: 60/080165
 / PRIOR FILING DATE: 1998-03-31
 / PRIOR APPLICATION NUMBER: 60/080194
 / PRIOR FILING DATE: 1998-03-31
 / PRIOR APPLICATION NUMBER: 60/080327
 / PRIOR FILING DATE: 1998-04-01
 / PRIOR APPLICATION NUMBER: 60/080328
 / PRIOR FILING DATE: 1998-04-01
 / PRIOR APPLICATION NUMBER: 60/080333
 / PRIOR FILING DATE: 1998-04-01
 / PRIOR APPLICATION NUMBER: 60/080334
 / PRIOR FILING DATE: 1998-04-01
 / PRIOR APPLICATION NUMBER: 60/081070
 / PRIOR FILING DATE: 1998-04-08
 / PRIOR APPLICATION NUMBER: 60/081049
 / PRIOR FILING DATE: 1998-04-08
 / PRIOR APPLICATION NUMBER: 60/081071
 / PRIOR FILING DATE: 1998-04-08
 / PRIOR APPLICATION NUMBER: 60/081195
 / PRIOR FILING DATE: 1998-04-08
 / PRIOR APPLICATION NUMBER: 60/081203
 / PRIOR FILING DATE: 1998-04-09
 / PRIOR APPLICATION NUMBER: 60/081229
 / PRIOR FILING DATE: 1998-04-09
 / PRIOR APPLICATION NUMBER: 60/081955
 / PRIOR FILING DATE: 1998-04-15
 / PRIOR APPLICATION NUMBER: 60/081817
 / PRIOR FILING DATE: 1998-04-15
 / PRIOR APPLICATION NUMBER: 60/081819
 / PRIOR FILING DATE: 1998-04-15
 / PRIOR APPLICATION NUMBER: 60/081952
 / PRIOR FILING DATE: 1998-04-15
 / PRIOR APPLICATION NUMBER: 60/081838
 / PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159

DB 12 GLLLLALL 19

RESULT 70

US-09-999-832A-85
; Sequence 85, Application US/0999832A

; Publication No. US20020192708A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James;

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2630P1C63

; CURRENT APPLICATION NUMBER: US/09/999,832A

; CURRENT FILING DATE: 2001-10-24

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077641

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077791

; PRIOR FILING DATE: 1998-03-12

1	PRIOR FILING DATE: 1998-04-22	
2	PRIOR APPLICATION NUMBER: 60/082700	
3	PRIOR FILING DATE: 1998-04-22	
4	PRIOR APPLICATION NUMBER: 60/082797	
5	PRIOR FILING DATE: 1998-04-22	
6	PRIOR APPLICATION NUMBER: 60/082796	
7	PRIOR FILING DATE: 1998-04-23	
8	PRIOR APPLICATION NUMBER: 60/083336	
9	PRIOR FILING DATE: 1998-04-27	
10	PRIOR APPLICATION NUMBER: 60/083322	
11	PRIOR FILING DATE: 1998-04-28	
12	PRIOR APPLICATION NUMBER: 60/083392	
13	PRIOR FILING DATE: 1998-04-29	
14	PRIOR APPLICATION NUMBER: 60/083495	
15	PRIOR FILING DATE: 1998-04-29	
16	PRIOR APPLICATION NUMBER: 60/083496	
17	PRIOR FILING DATE: 1998-04-29	
18	PRIOR APPLICATION NUMBER: 60/083499	
19	PRIOR FILING DATE: 1998-04-29	
20	PRIOR APPLICATION NUMBER: 60/083545	
21	PRIOR FILING DATE: 1998-04-29	
22	PRIOR APPLICATION NUMBER: 60/083554	
23	PRIOR FILING DATE: 1998-04-29	
24	PRIOR APPLICATION NUMBER: 60/083558	
25	PRIOR FILING DATE: 1998-04-29	
26	PRIOR APPLICATION NUMBER: 60/083559	
27	PRIOR FILING DATE: 1998-04-29	
28	PRIOR APPLICATION NUMBER: 60/083500	
29	PRIOR FILING DATE: 1998-04-29	
30	PRIOR APPLICATION NUMBER: 60/083742	
31	PRIOR FILING DATE: 1998-04-30	
32	PRIOR APPLICATION NUMBER: 60/084366	
33	PRIOR FILING DATE: 1998-05-05	
34	PRIOR APPLICATION NUMBER: 60/084414	
35	PRIOR FILING DATE: 1998-05-06	
36	PRIOR APPLICATION NUMBER: 60/084441	
37	PRIOR FILING DATE: 1998-05-06	
38	PRIOR APPLICATION NUMBER: 60/084637	
39	PRIOR FILING DATE: 1998-05-07	
40	PRIOR APPLICATION NUMBER: 60/084639	
41	PRIOR FILING DATE: 1998-05-07	
42	PRIOR APPLICATION NUMBER: 60/084640	
43	PRIOR FILING DATE: 1998-05-07	
44	PRIOR APPLICATION NUMBER: 60/084598	
45	PRIOR FILING DATE: 1998-05-07	
46	PRIOR APPLICATION NUMBER: 60/084600	
47	PRIOR FILING DATE: 1998-05-07	
48	PRIOR APPLICATION NUMBER: 60/084627	
49	PRIOR FILING DATE: 1998-05-07	
50	PRIOR APPLICATION NUMBER: 60/084643	
51	PRIOR FILING DATE: 1998-05-07	
52	PRIOR APPLICATION NUMBER: 60/085339	
53	PRIOR FILING DATE: 1998-05-13	
54	PRIOR APPLICATION NUMBER: 60/085338	
55	PRIOR FILING DATE: 1998-05-13	
56	PRIOR APPLICATION NUMBER: 60/085323	
57	PRIOR FILING DATE: 1998-05-13	
58	PRIOR APPLICATION NUMBER: 60/085579	
59	PRIOR FILING DATE: 1998-05-15	
60	PRIOR APPLICATION NUMBER: 60/085582	
61	PRIOR FILING DATE: 1998-05-15	
62	PRIOR APPLICATION NUMBER: 60/085700	
63	PRIOR FILING DATE: 1998-05-15	
64	PRIOR APPLICATION NUMBER: 60/085573	
65	PRIOR FILING DATE: 1998-05-15	
66	PRIOR APPLICATION NUMBER: 60/085704	
67	PRIOR FILING DATE: 1998-05-15	
68	PRIOR APPLICATION NUMBER: 60/085697	

Query Match 0.7% Score 8; DB 9; Length 67;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
 Db 12 GLLLLALL 19

RESULT 71

US-09-978-189-85
 ; Sequence 85, Application US/09978189
 ; Publication No. US20030004102A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gottschalk, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2630P1C7
 ; CURRENT APPLICATION NUMBER: US/09/978,189
 ; CURRENT FILING DATE: 2001-10-15
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079786
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079920
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/079923
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/080105
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080107
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080194
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080327
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080328
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080333
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080334
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/081070
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081049
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081071
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081195
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081203
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081229
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081955
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081817
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081819
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081952
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081838
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/082568
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082569
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082704
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082804
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082700
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082797
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082796
 ; PRIOR FILING DATE: 1998-04-23

; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085699
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
| | | | |
Db 12 GLLLLALL 19

RESULT 72

US-09-978-608A-85
; Sequence 85, Application US/09978608A
; Publication No. US20030045462A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kllavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C22
; CURRENT APPLICATION NUMBER: US/09/978,608A
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 85
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-608A-85

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
| | | | |
Db 12 GLLLLALL 19

RESULT 73

US-09-978-585A-85
; Sequence 85, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC15
; CURRENT APPLICATION NUMBER: US/09/978,585A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 85
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-585A-85

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 74
US-09-978-191A-85
; Sequence 85, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC4
; CURRENT APPLICATION NUMBER: US/09/978,191A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08

```


; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match

0.7%; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
 |||||
Db 12 GLLLLALL 19

RESULT 75

US-09-978-403A-85
; Sequence 85, Application US/09978403A
; Publication No. US20030050240A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC17
; CURRENT APPLICATION NUMBER: US/09/978,403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03

, PRIOR APPLICATION NUMBER: 60/065311
 , PRIOR FILING DATE: 1997-11-13
 , PRIOR APPLICATION NUMBER: 60/066364
 , PRIOR FILING DATE: 1997-11-21
 , PRIOR APPLICATION NUMBER: 60/077450
 , PRIOR FILING DATE: 1998-03-10
 , PRIOR APPLICATION NUMBER: 60/077632
 , PRIOR FILING DATE: 1998-03-11
 , PRIOR APPLICATION NUMBER: 60/077641
 , PRIOR FILING DATE: 1998-03-11
 , PRIOR APPLICATION NUMBER: 60/077649
 , PRIOR FILING DATE: 1998-03-11
 , PRIOR APPLICATION NUMBER: 60/077791
 , PRIOR FILING DATE: 1998-03-12
 , PRIOR APPLICATION NUMBER: 60/078004
 , PRIOR FILING DATE: 1998-03-13
 , PRIOR APPLICATION NUMBER: 60/078886
 , PRIOR FILING DATE: 1998-03-20
 , PRIOR APPLICATION NUMBER: 60/078936
 , PRIOR FILING DATE: 1998-03-20
 , PRIOR APPLICATION NUMBER: 60/078910
 , PRIOR FILING DATE: 1998-03-20
 , PRIOR APPLICATION NUMBER: 60/078939
 , PRIOR FILING DATE: 1998-03-20
 , PRIOR APPLICATION NUMBER: 60/079294
 , PRIOR FILING DATE: 1998-03-25
 , PRIOR APPLICATION NUMBER: 60/079656
 , PRIOR FILING DATE: 1998-03-26
 , PRIOR APPLICATION NUMBER: 60/079664
 , PRIOR FILING DATE: 1998-03-27
 , PRIOR APPLICATION NUMBER: 60/079689
 , PRIOR FILING DATE: 1998-03-27
 , PRIOR APPLICATION NUMBER: 60/079663
 , PRIOR FILING DATE: 1998-03-27
 , PRIOR APPLICATION NUMBER: 60/079728
 , PRIOR FILING DATE: 1998-03-27
 , PRIOR APPLICATION NUMBER: 60/079786
 , PRIOR FILING DATE: 1998-03-27
 , PRIOR APPLICATION NUMBER: 60/079920
 , PRIOR FILING DATE: 1998-03-30
 , PRIOR APPLICATION NUMBER: 60/079923
 , PRIOR FILING DATE: 1998-03-30
 , PRIOR APPLICATION NUMBER: 60/080105
 , PRIOR FILING DATE: 1998-03-31
 , PRIOR APPLICATION NUMBER: 60/080107
 , PRIOR FILING DATE: 1998-03-31
 , PRIOR APPLICATION NUMBER: 60/080165
 , PRIOR FILING DATE: 1998-03-31
 , PRIOR APPLICATION NUMBER: 60/080194
 , PRIOR FILING DATE: 1998-03-31
 , PRIOR APPLICATION NUMBER: 60/080327
 , PRIOR FILING DATE: 1998-04-01
 , PRIOR APPLICATION NUMBER: 60/080328
 , PRIOR FILING DATE: 1998-04-01
 , PRIOR APPLICATION NUMBER: 60/080333
 , PRIOR FILING DATE: 1998-04-01
 , PRIOR APPLICATION NUMBER: 60/080334
 , PRIOR FILING DATE: 1998-04-01
 , PRIOR APPLICATION NUMBER: 60/081070
 , PRIOR FILING DATE: 1998-04-08
 , PRIOR APPLICATION NUMBER: 60/081049
 , PRIOR FILING DATE: 1998-04-08
 , PRIOR APPLICATION NUMBER: 60/081071
 , PRIOR FILING DATE: 1998-04-08
 , PRIOR APPLICATION NUMBER: 60/081195
 , PRIOR FILING DATE: 1998-04-08
 , PRIOR APPLICATION NUMBER: 60/081203
 , PRIOR FILING DATE: 1998-04-09
 , PRIOR APPLICATION NUMBER: 60/081229
 , PRIOR FILING DATE: 1998-04-09
 , PRIOR APPLICATION NUMBER: 60/081955
 , PRIOR FILING DATE: 1998-04-15
 , PRIOR APPLICATION NUMBER: 60/081817

; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
|||||||
Db 12 GLLLLALL 19

Search completed: September 21, 2004, 14:09:50
Job time : 85 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:29:15 ; Search time 22 seconds

(without alignments)
2787.802 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVAWALSLWPGFT.....FRSARRREPGLDTPKYLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/protdata/2/iaa/5A-COMB.pep.*
- 2: /cgn2_6/protdata/2/iaa/5B-COMB.pep.*
- 3: /cgn2_6/protdata/2/iaa/6A-COMB.pep.*
- 4: /cgn2_6/protdata/2/iaa/6B-COMB.pep.*
- 5: /cgn2_6/protdata/2/iaa/PCTUS-COMB.pep.*
- 6: /cgn2_6/protdata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	1.8	158	4	US-09-312-283C-391
2	17	1.4	85	4	US-09-312-283C-390
3	15	1.3	214	4	US-10-061-658-5
4	15	1.3	214	4	US-10-061-658-6
5	15	1.3	214	4	US-10-061-658-9
6	14	1.2	1183	4	US-09-532-310B-5
7	11	0.9	22	6	5196511-9
8	11	0.9	1183	4	US-09-532-310B-6
9	9	0.8	9	1	US-08-462-128-34
10	9	0.8	9	1	US-08-463-180-34
11	8	0.7	17	1	US-08-173-497-20
12	8	0.7	17	1	US-08-286-889-20
13	8	0.7	17	1	US-08-485-618-20
14	8	0.7	17	1	US-08-362-652-20
15	8	0.7	17	2	US-08-605-672-20
16	8	0.7	17	2	US-08-482-293A-20
17	8	0.7	17	2	US-08-943-363-20
18	8	0.7	17	3	US-09-193-043-20
19	8	0.7	17	4	US-09-688-307A-20
20	8	0.7	17	4	US-09-350-259-20
21	8	0.7	76	1	US-08-848-252-4
22	8	0.7	76	3	US-09-083-521-5
23	8	0.7	80	1	US-08-848-252-2
24	8	0.7	95	1	US-07-626-618A-2
25	8	0.7	95	1	US-07-928-611-2
26	8	0.7	95	1	US-08-333-977-2
27	8	0.7	95	2	US-08-487-811A-2

28	8	0.7	95	3	US-09-060-694-2	Sequence 2, Appli
29	8	0.7	95	4	US-09-378-074-2	Sequence 2, Appli
30	8	0.7	95	5	PCT-US93-07370-2	Sequence 2, Appli
31	8	0.7	97	4	US-08-973-544-4	Sequence 4, Appli
32	8	0.7	104	4	US-08-973-544-3	Sequence 3, Appli
33	8	0.7	112	4	US-09-489-039A-10236	Sequence 10236, A
34	8	0.7	179	4	US-09-419-568F-28	Sequence 28, Appl
35	8	0.7	179	4	US-09-354-243B-28	Sequence 28, Appl
36	8	0.7	179	4	US-09-870-574-2	Sequence 2, Appli
37	8	0.7	183	4	US-09-252-991A-25905	Sequence 25905, A
38	8	0.7	234	3	US-09-040-483-5	Sequence 5, Appli
39	8	0.7	234	4	US-08-740-036-5	Sequence 5, Appli
40	8	0.7	315	1	US-08-118-270-28	Sequence 28, Appl
41	8	0.7	315	5	PCT-US93-08528-28	Sequence 28, Appl
42	8	0.7	345	4	US-09-489-039A-7417	Sequence 7417, Ap
43	8	0.7	374	3	US-09-046-736-4	Sequence 4, Appli
44	8	0.7	387	1	US-07-626-618A-17	Sequence 17, Appl
45	8	0.7	387	1	US-08-086-439C-3	Sequence 3, Appli
46	8	0.7	387	1	US-08-056-051-2	Sequence 2, Appli
47	8	0.7	387	1	US-07-928-611-18	Sequence 18, Appl
48	8	0.7	387	1	US-08-333-977-17	Sequence 17, Appl
49	8	0.7	387	1	US-08-449-153-1	Sequence 1, Appli
50	8	0.7	387	1	US-08-434-877-3	Sequence 3, Appli
51	8	0.7	387	1	US-08-449-851-1	Sequence 1, Appli
52	8	0.7	387	2	US-08-833-703-1	Sequence 1, Appli
53	8	0.7	387	3	US-08-487-811A-18	Sequence 18, Appl
54	8	0.7	387	3	US-08-475-742-4	Sequence 4, Appli
55	8	0.7	387	3	US-09-060-694-18	Sequence 18, Appl
56	8	0.7	387	4	US-09-378-074-18	Sequence 18, Appl
57	8	0.7	387	4	US-08-261-293-4	Sequence 4, Appli
58	8	0.7	387	5	PCT-US93-07370-18	Sequence 18, Appl
59	8	0.7	419	1	US-08-056-051-4	Sequence 4, Appli
60	8	0.7	419	1	US-07-928-611-20	Sequence 20, Appl
61	8	0.7	419	2	US-08-487-811A-20	Sequence 20, Appl
62	8	0.7	419	3	US-09-060-694-20	Sequence 20, Appl
63	8	0.7	419	4	US-09-378-074-20	Sequence 20, Appl
64	8	0.7	419	5	PCT-US93-07370-20	Sequence 20, Appl
65	8	0.7	453	4	US-09-252-991A-17394	Sequence 17394, A
66	8	0.7	467	1	US-08-056-051-6	Sequence 6, Appli
67	8	0.7	467	1	US-07-928-611-22	Sequence 22, Appl
68	8	0.7	467	2	US-08-487-811A-22	Sequence 22, Appl
69	8	0.7	467	3	US-09-046-736-2	Sequence 2, Appli
70	8	0.7	467	3	US-09-060-694-22	Sequence 22, Appl
71	8	0.7	467	4	US-09-378-074-22	Sequence 22, Appl
72	8	0.7	467	5	PCT-US93-07370-22	Sequence 22, Appl
73	8	0.7	494	1	US-08-485-618-103	Sequence 103, App
74	8	0.7	494	2	US-08-605-672-103	Sequence 103, App
75	8	0.7	494	2	US-08-482-293A-103	Sequence 103, App
76	8	0.7	494	2	US-08-943-363-103	Sequence 103, App
77	8	0.7	494	3	US-09-193-043-103	Sequence 103, App
78	8	0.7	494	4	US-09-688-307A-103	Sequence 103, App
79	8	0.7	494	4	US-09-350-259-103	Sequence 103, App
80	8	0.7	605	2	US-08-472-666-1	Sequence 1, Appli
81	8	0.7	605	5	PCT-US96-07615-1	Sequence 1, Appli
82	8	0.7	628	4	US-09-252-991A-32727	Sequence 32727, A
83	8	0.7	877	4	US-09-252-991A-25547	Sequence 25547, A
84	8	0.7	1151	1	US-08-286-889-37	Sequence 37, Appl
85	8	0.7	1151	1	US-08-485-618-37	Sequence 37, Appl
86	8	0.7	1151	1	US-08-362-652-37	Sequence 37, Appl
87	8	0.7	1151	2	US-08-605-672-37	Sequence 37, Appl
88	8	0.7	1151	2	US-08-482-293A-37	Sequence 37, Appl
89	8	0.7	1151	2	US-08-943-363-37	Sequence 37, Appl
90	8	0.7	1151	3	US-09-193-043-37	Sequence 37, Appl
91	8	0.7	1151	4	US-09-688-307A-37	Sequence 37, Appl
92	8	0.7	1151	4	US-09-350-259-37	Sequence 37, Appl
93	8	0.7	1152	2	US-08-476-062A-43	Sequence 43, Appl
94	8	0.7	1152	5	PCT-US96-01314-43	Sequence 43, Appl
95	8	0.7	1152	6	5424399-2	Patent No. 5424399
96	8	0.7	1153	1	US-08-173-497-3	Sequence 3, Appli
97	8	0.7	1153	1	US-08-286-889-3	Sequence 3, Appli
98	8	0.7	1153	1	US-08-485-618-3	Sequence 3, Appli
99	8	0.7	1153	1	US-08-362-652-3	Sequence 3, Appli
100	8	0.7	1153	2	US-08-605-672-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-312-283C-391

; Sequence 391, Application US/09312283C
; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 391

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-391

Query Match 1.8%; Score 21; DB 4; Length 158;

Best Local Similarity 100.0%; Pred. No. 1e-11; Indels 0;

Matches 21; Conservative 0; Mismatches 0; Gaps 0;

Qy 636 RPVVQINASLHPEPSKINIFH 656

Db 57 RPVVQINASLHPEPSKINIFH 77

RESULT 2

US-09-312-283C-390

; Sequence 390, Application US/09312283C

; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 390

; LENGTH: 85

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-390

Query Match 1.4%; Score 17; DB 4; Length 85;

Best Local Similarity 100.0%; Pred. No. 3.7e-08; Indels 0;

Matches 17; Conservative 0; Mismatches 0; Gaps 0;

Qy 528 NGTLKDSHSYQNARFGS 544

Db 28 NGTLKDSHSYQNARFGS 44

RESULT 3

US-10-061-658-5

; Sequence 5, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Koteliarsky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Rat

US-10-061-658-5

Query Match 1.3%; Score 15; DB 4; Length 214;

Best Local Similarity 100.0%; Pred. No. 6.6e-06; Indels 0;

Matches 15; Conservative 0; Mismatches 0; Gaps 0;

Qy 164 DIVVLGDGNSIYFW 178

Db 22 DIVVLGDGNSIYFW 36

RESULT 4

US-10-061-658-6

; Sequence 6, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Koteliarsky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-061-658-6

Query Match 1.3%; Score 15; DB 4; Length 214;

Best Local Similarity 100.0%; Pred. No. 6.6e-06; Indels 0;

Matches 15; Conservative 0; Mismatches 0; Gaps 0;

Qy 164 DIVVLGDGNSIYFW 178

Db 22 DIVVLGDGNSIYFW 36

RESULT 5

US-10-061-658-9

; Sequence 9, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Koteliarsky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

```
FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

Query Match      1.3%; Score 15; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGGSNSIYP 178
22 DIVVLGGSNSIYP 36
Db

RESULT 6
US-09-532-310B-5
; Sequence 5, Application US/09532310B
; Patent No. 6594276
; GENERAL INFORMATION:
; APPLICANT: Senger, Donald R
; Detmar, Michael
; Claffey, Kevin P
; TITLE OF INVENTION: Method for inhibiting tumor
; angiogenesis in a living subject
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David Prashker, Esq.
; STREET: P.O. Box 5387
; CITY: Magnolia
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 01930
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
; COMPUTER: Dell PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Microsoft Word version 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/532,310B
; FILING DATE: 22-Mar-2000
; CLASSIFICATION: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: David Prashker, Esq.
; REGISTRATION NUMBER: 29,693
; REFERENCE/DOCKET NUMBER: BIS-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (978) 525-3794
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-532-310B-5

Query Match      1.2%; Score 14; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGGSNSIYP 177
|||||
```

```
Db 174 DIVVLGGSNSIYP 187

RESULT 7
US-09-532-310B-6
; Sequence 5, Application US/09532310B
; Patent No. 6594276
; GENERAL INFORMATION:
; APPLICANT: Senger, Donald R
; Detmar, Michael
; Claffey, Kevin P
; TITLE OF INVENTION: Method for inhibiting tumor
; angiogenesis in a living subject
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David Prashker, Esq.
; STREET: P.O. Box 5387
; CITY: Magnolia
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 01930
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
; COMPUTER: Dell PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Microsoft Word version 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/532,310B
; FILING DATE: 22-Mar-2000
; CLASSIFICATION: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: David Prashker, Esq.
; REGISTRATION NUMBER: 29,693
; REFERENCE/DOCKET NUMBER: BIS-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (978) 525-3794
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-532-310B-6

Query Match      0.9%; Score 11; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 495 TDVLLVGAPMY 505
 |||||
 Db 506 TDVLLVGAPMY 516

RESULT 9

US-08-462-128-34
 ; Sequence 34, Application US/08462128
 ; Patent No. 5686059
 ; GENERAL INFORMATION:
 ; APPLICANT: Goetinck, Paul F.
 ; APPLICANT: Tondravi, Mehrdad
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lahive & Cockfield
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/462,128
 FILING DATE: 15-JAN-1993
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/006,096
 FILING DATE: 15-JAN-1993
 APPLICATION NUMBER: US 07/866,403
 FILING DATE: 10-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: MGP-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-462-128-34

Query Match 0.8%; Score 9; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 DIVVLDS 172
 |||||
 Db 1 DIVVLDS 9

RESULT 10

US-08-463-180-34
 ; Sequence 34, Application US/08463180
 ; Patent No. 5741670
 ; GENERAL INFORMATION:
 ; APPLICANT: Goetinck, Paul F.
 ; APPLICANT: Tondravi, Mehrdad
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lahive & Cockfield
 ; STREET: 60 State Street

CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/463,180
 FILING DATE: 15-JAN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/006,096
 FILING DATE: 15-JAN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/866,403
 FILING DATE: 10-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: MGP-005DV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-463-180-34

Query Match 0.8%; Score 9; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 DIVVLDS 172
 |||||
 Db 1 DIVVLDS 9

RESULT 11

US-08-173-497-20
 ; Sequence 20, Application US/08173497
 ; Patent No. 5437958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, W. Michael
 ; APPLICANT: Van Der Vieren, Monica
 ; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
 ; TITLE OF INVENTION: Subunit
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 S. Wacker Drive, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/173,497
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5437958and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27866/31363

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-173-497-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 12
US-08-286-889-20
Sequence 20, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-286-889-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 13
US-08-485-618-20
Sequence 20, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-485-618-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 14
US-08-362-652-20
Sequence 20, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE: 5-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-652-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 15
US-08-605-672-20
Sequence 20, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-605-672-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 16
US-08-482-293A-20
Sequence 20, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLACULE TYPE: peptide
US-08-482-293A-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 17

US-08-943-363-20
; Sequence 20, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLACULE TYPE: peptide
US-08-943-363-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 18
US-09-193-043-20
; Sequence 20, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-193-043-20

Query Match 0.7%; Score 8; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 19

US-09-688-307A-20
; Sequence 20, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 13
; OTHER INFORMATION: Xaa = any or unknown amino acid
US-09-688-307A-20

Query Match 0.7%; Score 8; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
|||||||
DB 1 LVVGAPLE 8

RESULT 20

US-09-350-259-20
; Sequence 20, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-350-259-20

Query Match 0.7%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
|||||||
DB 1 LVVGAPLE 8

RESULT 21

US-08-848-252-4
; Sequence 4, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/06/538,052
; FILING DATE:

APPLICATION NUMBER: US 08/151,672
FILING DATE: 15-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mcdiarmid, Shona S.
REGISTRATION NUMBER: P-38,798
REFERENCE/DOCKET NUMBER: 3158-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
TELEX: 06-23115
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-848-252-4

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLALL 1159
|||||||
DB 12 GLLLALL 19

RESULT 22

US-09-083-521-5
; Sequence 5, Application US/09083521
; Patent No. 6048970
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1216498
US-09-083-521-5

Query Match 0.7%; Score 8; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159
Db 12 GLLALL 19

RESULT 23
US-08-848-252-2
; Sequence 2, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,052
; FILING DATE:
; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MCDiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-848-252-2

Query Match 0.7%; Score 8; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159
Db 12 GLLALL 19

RESULT 24
US-07-626-618A-2
; Sequence 2, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/626,618A
FILING DATE: 7 DEC 1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5422265nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-626-618A-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 25
US-07-928-611-2
; Sequence 2, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611
FILING DATE: 19920810
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5569601nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 95 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-928-611-2

Query Match 0.7%; Score 8; DB 1; Length 95;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 Db 81 LLLALLVL 88

RESULT 26

US-08-333-977-2
 ; Sequence 2, Application US/08333977
 ; Patent No. 5594108

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.
 APPLICANT: Civelli, Olivier
 TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.
 STREET: 10 South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/333,977
 FILING DATE: 03-NOV-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/626,618
 FILING DATE: 7 DEC 1990

ATTORNEY/AGENT INFORMATION:

NAME: No. 5594108nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90,1092

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 95 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-333-977-2

Query Match 0.7%; Score 8; DB 1; Length 95;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 Db 81 LLLALLVL 88

RESULT 27

US-08-487-811A-2
 ; Sequence 2, Application US/08487811A
 ; Patent No. 5883226

GENERAL INFORMATION:

APPLICANT: Civelli, Olivier
 APPLICANT: Van Tol, Hubert H.M.
 TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 South Wacker Drive
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A
 FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 95 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-487-811A-2

Query Match 0.7%; Score 8; DB 2; Length 95;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 Db 81 LLLALLVL 88

RESULT 28

US-09-060-694-2
 ; Sequence 2, Application US/09060694
 ; Patent No. 6203998

GENERAL INFORMATION:

APPLICANT: Civelli, Olivier
 APPLICANT: Van Tol, Hubert H.M.
 TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 South Wacker Drive
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060,694
 FILING DATE: 15-APR-1998

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 6203998nan, Kevin E
 REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-MM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-694-2

Query Match 0.7%; Score 8; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||||
DB 81 LLLALLVL 88

RESULT 29
US-09-378-074-2
Sequence 2, Application US/09378074
Patent No. 6437114
GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.
Civelli, Olivier
TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/378,074
FILING DATE: 20-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/928,611
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6437114han, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-378-074-2

Query Match 0.7%; Score 8; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||||

DB 81 LLLALLVL 88
RESULT 30
PCT-US93-07370-2
Sequence 2, Application PC/TUS9307370
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07370
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-07370-2

Query Match 0.7%; Score 8; DB 5; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||||
DB 81 LLLALLVL 88

RESULT 31
US-08-973-544-4
Sequence 4, Application US/08973544
Patent No. 6338950
GENERAL INFORMATION:
APPLICANT: WEISS, Elisabeth
TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,544
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT EP 96/02663
FILING DATE: 20-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95109511.6
FILING DATE: 20-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95112201.9
FILING DATE: 03-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kitts, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P8341-7073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000

TELEFAX: (202) 638-4810
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 97 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-973-544-4

Query Match 0.7%; Score 8; DB 4; Length 97;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157
 Db 16 LGGLLLLA 23
 |||||

RESULT 32
 US-08-973-544-3
 ; Sequence 3, Application US/08973544
 ; Patent No. 6338950
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, Elisabeth
 ; TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIKALDO, MARMELESTEIN, MURRAY & ORAM LLP
 ; STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-5701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/973,544
 ; FILING DATE: 18-DEC-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT EP 96/02663
 ; FILING DATE: 20-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 95109511.6
 ; FILING DATE: 20-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 95112201.9
 ; FILING DATE: 03-AUG-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitts, Monica Chin
 ; REGISTRATION NUMBER: 36,105
 ; REFERENCE/DOCKET NUMBER: P8341-7073
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 638-5000
 ; TELEFAX: (202) 638-4810
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 104 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-973-544-3

Query Match 0.7%; Score 8; DB 4; Length 104;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157
 |||||

Db 16 LGGLLLLA 23

RESULT 33
 US-09-489-039A-10236
 ; Sequence 10236, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 10236
 ; LENGTH: 112
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-10236

Query Match 0.7%; Score 8; DB 4; Length 112;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 Db 13 LLLALLV 20
 |||||

RESULT 34
 US-09-419-568F-28
 ; Sequence 28, Application US/09419568F
 ; Patent No. 6331613
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumoutier, Laure
 ; APPLICANT: Louhed, Jamila
 ; APPLICANT: Renauld, Jean-Christophe
 ; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac
 ; TITLE OF INVENTION: (TIFs) The Proteins Encoded, and Uses Thereof
 ; FILE REFERENCE: LUD 5543.2
 ; CURRENT APPLICATION NUMBER: US/09/419,568F
 ; CURRENT FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: US09/354,243
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: US09/178,973
 ; PRIOR FILING DATE: 1998-10-26
 ; NUMBER OF SEQ ID NOS: 29
 ; SEQ ID NO 28
 ; LENGTH: 179
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; US-09-419-568F-28

Query Match 0.7%; Score 8; DB 4; Length 179;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 Db 21 LLLALLV 28
 |||||

RESULT 35
 US-09-354-243B-28
 ; Sequence 28, Application US/09354243B
 ; Patent No. 6359117
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumoutier, Laure
 ; APPLICANT: Louhed, Jamila


```

; APPLICANT: Renault, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
; TITLE OF INVENTION: (RIFs)
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof
; FILE REFERENCE: LUD 5543.1
; CURRENT APPLICATION NUMBER: US/09/354,243B
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-28

Query Match
Best Local Similarity 100.0%; Pred. No. 24; Length 179;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
Db 21 LLLLLLV 28

RESULT 36
US-09-870-574-2
; Sequence 2, Application US/09870574
; Patent No. 6551799
; GENERAL INFORMATION:
; APPLICANT: Gurney, Austin L.
; APPLICANT: Aggarwal, Sudeepa
; APPLICANT: Xie, Ming-Hong
; APPLICANT: Maruoka, Ellen M.
; APPLICANT: Foster, Jessica S.
; APPLICANT: Goddard, Audrey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: INTERLEUKIN-22 POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THE SAME AND METHODS FOR THE TREATMENT OF PANCREATIC DISORDERS
; FILE REFERENCE: P2806-1(US)
; CURRENT APPLICATION NUMBER: US/09/870,574
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/169,495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-870-574-2

Query Match
Best Local Similarity 100.0%; Pred. No. 24; Length 179;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
Db 21 LLLLLLV 28

RESULT 37
US-09-252-991A-25905
; Sequence 25905, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25905
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25905

Query Match
Best Local Similarity 100.0%; Pred. No. 25; Length 183;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1171 SARRRREP 1178
Db 21 SARRRREP 28

RESULT 38
US-09-040-483-5
; Sequence 5, Application US/09040483
; Patent No. 6143867
; GENERAL INFORMATION:
; APPLICANT: Aketdiom, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; TITLE OF INVENTION: DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Icyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,483
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/740,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
US-09-040-483-5

Query Match
Best Local Similarity 100.0%; Pred. No. 31; Length 234;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1153 LLLALLV 1160
Db 4 LLLALLV 11

RESULT 39
US-08-740-036-5
; Sequence 5, Application US/08740036
; Patent No. 6492507
; GENERAL INFORMATION:
; APPLICANT: Aketblom, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; TITLE OF INVENTION: DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,036
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
; US-08-740-036-5

Query Match 0.7%; Score 8; DB 4; Length 234;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
Db 4 LLLALLV 11

RESULT 40
US-08-118-270-28
; Sequence 28, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197

```

```

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-28

Query Match 0.7%; Score 8; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1161
Db 48 LLLALLV 55

RESULT 41
PCT-US93-08528-28
; Sequence 28, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197

```

```

; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-28

Query Match 0.7%; Score 8; DB 5; Length 315;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 48 LLLALLVL 55

RESULT 42
US-09-489-039A-7417
; Sequence 7417, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7417
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7417

Query Match 0.7%; Score 8; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158
DB 76 GGLLLAL 83

RESULT 43
US-09-046-736-4
; Sequence 4, Application US/09046736
; Patent No. 6090582
; GENERAL INFORMATION:
; APPLICANT: KIKUY, KRISTINE
; APPLICANT: ERICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,736
; FILING DATE: 24-MAR-1998

```

```

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,885
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-736-4

Query Match 0.7%; Score 8; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GVLLGAVG 386
DB 257 GVLLGAVG 264

RESULT 44
US-07-626-618A-17
; Sequence 17, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/626,618A
; FILING DATE: 7 DEC 1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5422265pan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-07-626-618A-17

Query Match 0.7%; Score 8; DB 1; Length 387;

```



```
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-928-611-18

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 48
US-08-333-977-17
; Sequence 17, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; US-08-333-977-17

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 49
US-08-449-153-1
; Sequence 1, Application US/08449153
; Patent No. 5686255
; GENERAL INFORMATION:
; APPLICANT: Deth, Richard C.
; TITLE OF INVENTION: Compositions and Methods for Diagnosing
; TITLE OF INVENTION: Schizophrenia
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,153
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: RCD95-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-8240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-449-153-1

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 50
US-08-434-877-3
; Sequence 3, Application US/08434877
; Patent No. 5721132
; GENERAL INFORMATION:
; APPLICANT: Chio, Christopher L.
; APPLICANT: Huff, Rita M.
; TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
; TITLE OF INVENTION: Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Upjohn Company, Corp. Intellectual
; ADDRESS: Property Law
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette (DS,HD)
; COMPUTER: Gateway 2000, P5-90
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,877
```

;; FILING DATE: 1 July 1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dainley Jr., James D.
;; REGISTRATION NUMBER: 33,673
;; REFERENCE/DOCKET NUMBER: 4700 DVI
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 616-385-5210
;; TELEFAX: 616-385-6897
;; TELEX: 224401
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 387 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-434-877-3

Query Match 0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 51
US-08-449-651-1
; Sequence 1, Application US/08449651
; Patent No. 5738998
; GENERAL INFORMATION:
; APPLICANT: Deth, Richard C.
; TITLE OF INVENTION: Compositions and Methods for Diagnosing
; TITLE OF INVENTION: Schizophrenia
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,651
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: RCD95-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-449-651-1

Query Match 0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 52
US-08-487-811A-18
; Sequence 18, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Robert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nat, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-18

Query Match 0.7%; Score 8; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 53
US-08-833-703-1
; Sequence 1, Application US/08833703
; Patent No. 6080549
; GENERAL INFORMATION:
; APPLICANT: DETH, RICHARD
; TITLE OF INVENTION: METHODS AND MATERIALS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF SCHIZOPHRENIA AND RELATED DISORDERS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,703
; FILING DATE: 08-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Ph.D., Holliday C
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: NU-431XX
; TELEPHONE: 617-542 2290
; TELEFAX: 617-451 0313
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-833-703-1

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
DB 81 LLLALLVL 88

RESULT 54
US-08-475-742-4
; Sequence 4, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L
; APPLICANT: Todd, Richard D
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
; FILE REFERENCE: WU 102 CON DIV
; CURRENT APPLICATION NUMBER: US/08/475,742
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: US 08/261,293
; EARLIER FILING DATE: 1994-06-16
; EARLIER APPLICATION NUMBER: US 08/014,013
; EARLIER FILING DATE: 1993-01-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(387)
; OTHER INFORMATION: Human D4 Receptor Protein
; US-08-475-742-4

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
DB 81 LLLALLVL 88

RESULT 55
US-09-060-694-18
; Sequence 18, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-WM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-694-18

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
DB 81 LLLALLVL 88

RESULT 56
US-09-378-074-18
; Sequence 18, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999
; CLASSIFICATION: <Unknown>
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/928,611
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 643711nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-378-074-18

Query Match 0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 57
US-08-261-293-4
Sequence 4, Application US/08261293
Patent No. 6486310
GENERAL INFORMATION:
APPLICANT: O'Malley, Karen L.
APPLICANT: Todd, Richard D.
TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4
NUMBER OF INVENTION: Receptor
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,293
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014,013
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WU 102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6524
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..387
OTHER INFORMATION: /note= "Human D4 Receptor Protein"
US-08-261-293-4

Query Match 0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 58
PCT-US93-07370-18
Sequence 18, Application PC/TUS9307370
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07370
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-07370-18

Query Match 0.7%; Score 8; DB 5; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 59
US-08-056-051-4
Sequence 4, Application US/08056051
Patent No. 5516683
GENERAL INFORMATION:
APPLICANT: Grandy, David K
APPLICANT: Bunzow, James R
APPLICANT: Civeilli, Olivier
APPLICANT: Van Tol, Hubert H.-M.
TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/056,051
;; FILING DATE: 19930429
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5516683nan, Kevin E
;; REGISTRATION NUMBER: 35,303
;; REFERENCE/DOCKET NUMBER: 90,1092-C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; TELEX: 910-221-5317
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 419 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-056-051-4

Query Match 0.7%; Score 8; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 60
US-07-928-611-20
; Sequence 20, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-20

Query Match 0.7%; Score 8; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 61
US-08-487-811A-20
; Sequence 20, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-20

Query Match 0.7%; Score 8; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 62
US-09-060-694-20
; Sequence 20, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/060,694
 ; FILING DATE: 15-APR-1998
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 6203998nan, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 90,1092-MM
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-913-0001
 ; TELEFAX: 312-913-0002
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 419 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-060-694-20

Query Match 0.7%; Score 8; DB 3; Length 419;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 DB 81 LLLALLVL 88

RESULT 63
 US-09-378-074-20
 ; Sequence 20, Application US/09378074
 ; Patent No. 6437114
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Tol, Hubert H.M.
 ; Civeilli, Olivier
 ; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Allegretti & Witcoff, Ltd.
 ; STREET: 10 South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/378,074
 ; FILING DATE: 20-Aug-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/928,611
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 6437114nan, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 90,1092-B
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 810-221-8317
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 419 amino acids
 ; TYPE: amino acid

; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
 ; US-09-378-074-20

Query Match 0.7%; Score 8; DB 4; Length 419;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 DB 81 LLLALLVL 88

RESULT 64
 PCT-US93-07370-20
 ; Sequence 20, Application PC/TUS9307370
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
 ; NUMBER OF SEQUENCES: 22
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/07370
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 419 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US93-07370-20

Query Match 0.7%; Score 8; DB 5; Length 419;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
 |||||
 DB 81 LLLALLVL 88

RESULT 65
 US-09-252-991A-17394
 ; Sequence 17394, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; SOFTWARE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17394
 ; LENGTH: 453
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-17394

Query Match 0.7%; Score 8; DB 4; Length 453;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157
 |||||

Db 426 LGGLLLLA 433

RESULT 66

US-08-056-051-6

Sequence 6, Application US/08056051

Patent No. 5516683

GENERAL INFORMATION:

APPLICANT: Grandy, David K

APPLICANT: Bunzow, James R

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.-M.

TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,051

FILING DATE: 19930429

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5516683nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-056-051-6

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 67

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 68

US-08-487-811A-22

Sequence 22, Application US/08487811A

Patent No. 5883226

GENERAL INFORMATION:

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.M.

TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/928,611

FILING DATE: 19920810

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5569601nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-9317

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-928-611-22

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Grandy, David K

APPLICANT: Bunzow, James R

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.-M.

TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,051

FILING DATE: 19930429

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5516683nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-056-051-6

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 67

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 68

US-08-487-811A-22

Sequence 22, Application US/08487811A

Patent No. 5883226

GENERAL INFORMATION:

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.M.

TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Query Match 0.7%; Score 8; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 69

US-09-046-736-2
; Sequence 2, Application US/09046736
; Patent No. 6090582
; GENERAL INFORMATION:
; APPLICANT: KIRBY, KRISTINE
; APPLICANT: BRICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,736
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,885
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-736-2

Query Match 0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GVLGAVG 386
DB 350 GVLGAVG 357

RESULT 70

US-09-060-694-22
; Sequence 22, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998

CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:

INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-060-694-22

Query Match 0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 71

US-09-378-074-22
; Sequence 22, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999

CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6437114nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,559
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-103

Query Match 0.7%; Score 8; DB 1; Length 494;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLL 1158
Db 442 GGLLLLL 449

RESULT 74
US-08-605-672-103
Sequence 103, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
ADDRESS: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:

```

NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-103

Query Match 0.7%; Score 8; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1151 GGLLLAL 1158
|||
Db 442 GGLLLAL 449

RESULT 75

US-08-482-293A-103
Sequence 103, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vliet, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,437
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-103

Query Match 0.7%; Score 8; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1151 GGLLLAL 1158
|||
Db 442 GGLLLAL 449

Search completed: September 21, 2004, 13:42:57
Job time : 24 secs